

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 15, 2005, 02:48:24 ; Search time 4914 Seconds  
(without alignments)

2987.776 Million cell updates/sec

Title: US-10-715-665-7\_COPY\_1\_303

Perfect score: 1746

Sequence: 1 MDAMKRGCCVLLCGAVFV.....AACNWTGRGRCLEDRDRSE 303

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DSV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10715665/runat\_13042005.165822.28138/app.query.fasta\_1.455  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-FQ=US10715665 @CGN 1.1 3731 @runat\_13042005.165822.28138 -NCFU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.hg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1746	100.0	5882	6 AX154501	AX154501 Sequence
2	1620.5	92.8	1207	6 AR004336	AR004336 Sequence
3	1620.5	92.8	1207	6 AR006827	AR006827 Sequence
4	1620.5	92.8	1210	6 AR031233	AR031233 Sequence

5	1620.5	92.8	1210	6	AR145049	Sequence
6	1620.5	92.8	1989	6	AR179260	Sequence
7	1620.5	92.8	1989	6	BD132859	Intracell
8	1620.5	92.8	1989	6	AR282014	Sequence
9	1620.5	92.8	8316	6	AR118703	Sequence
10	1620.5	92.8	8987	6	AR118728	Sequence
11	1620.5	92.8	9185	6	AR118722	Sequence
12	1620.5	92.8	9185	6	AR118723	Sequence
13	1620.5	92.8	9185	6	108294	Sequence 1
14	1620.5	92.8	9185	6	BD091382	HCV culti
15	1620.5	92.8	9379	6	AR166930	Sequence
16	1620.5	92.8	9379	6	AR301300	Sequence
17	1620.5	92.8	9401	6	AR176483	Sequence
18	1620.5	92.8	9401	6	E66593	Hepatitis C
19	1620.5	92.8	9401	6	171894	Sequence 9
20	1620.5	92.8	9401	6	181885	Sequence 9
21	1620.5	92.8	9401	6	BD080334	Hepatitis
22	1620.5	92.8	9401	14	HPCFLYPRE	M62321 Hepatitis C
23	1620.5	92.8	9609	12	AF387805	AF387805 Synthetic
24	1620.5	92.8	9609	12	AF387808	AF387808 Synthetic
25	1620.5	92.8	9646	12	AF387806	AF387806 Synthetic
26	1620.5	92.8	9693	12	AF387807	AF387807 Synthetic
27	1615.5	92.5	9379	6	AR118747	Sequence
28	1608.5	92.1	9618	14	AF271632	AF271632 Hepatitis
29	1530.5	87.7	7178	6	CQ785592	Sequence
30	1530.5	87.7	7178	6	CQ795334	Sequence
31	1530.5	87.7	7178	6	CQ867238	Sequence
32	1526	87.4	5125	6	AR068510	Sequence
33	1526	87.4	5125	6	165457	Sequence 11
34	1526	87.4	5323	6	AR068509	Sequence
35	1526	87.4	5323	6	165456	Sequence 9
36	1523.5	87.3	2610	14	HPCST77	M62381 Hepatitis C
37	1521.5	87.1	7551	6	CQ795336	Sequence
38	1521.5	87.1	7551	6	CQ867240	Sequence
39	1521.5	87.1	7754	6	CQ785590	Sequence
40	1521.5	87.1	7754	6	CQ795332	Sequence
41	1521.5	87.1	7754	6	CQ867236	Sequence
42	1521.5	87.1	9401	6	AR030378	Sequence
43	1521.5	87.1	9416	6	AR110845	Sequence
44	1521.5	87.1	9416	6	BD069999	Functiona
45	1521.5	87.1	9622	6	AX663428	Sequence

#### ALIGNMENTS

RESULT 1	AX154501	AX154501	Sequence 6 from Patent WO0138358.	DNA	linear	PAT 22-JUN-2001
LOCUS	AX154501	Sequence 6 from Patent WO0138358.	5882 bp	DNA	linear	PAT 22-JUN-2001
DEFINITION	AX154501	Sequence 6 from Patent WO0138358.	5882 bp	DNA	linear	PAT 22-JUN-2001
ACCESSION	AX154501	Sequence 6 from Patent WO0138358.	5882 bp	DNA	linear	PAT 22-JUN-2001
VERSION	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
KEYWORDS	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
SOURCE	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
ORGANISM	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
REFERENCE	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
AUTHORS	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
TITLE	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
JOURNAL	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
FEATURES	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
source	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
CDS	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
Location/Qualifiers	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
1. .5882	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/organism="synthetic construct"	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/mol_type="unassigned DNA"	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/db_xref="taxon:32630"	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/note="plasmid pCMV-II-E2661-sag"	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
1992. .3587	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/note="unnamed protein product"	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/codon_start=1	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/transl_table=11	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/protein_id="CAC42602.1"	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/db_xref="GI:14536112"	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001
/translation="MDAMKRGCCVLLCGAVFVSPSPASETHVTGSGAGHTVSGFVSL"	AX154501.1	GI:14536111	GI:14536111	DNA	linear	PAT 22-JUN-2001

LAPGAKQNVQLINTNGSWHLNLTALNCNDLSLNTGLAGLFVYHHKNSGCPERLASCR  
PUDFDDQMGPISTYANGSGPDORPCWHPKPRCGIVPAKSVCGVYCFPTSPVVVGT  
TDRSGAPTYSGENDTDVFLNNTPPPLGNWFGCTWMNSTFTKVGCGAPPCVIGAGN  
NTLHGPTDCFRXHPDAPTSYRCGSGPMITPRCLVDYPRYLWHYPCTINTYIFKIRMYVG  
GVHEHLEAACNWRGRCLEDRDRSEIDMENITSGFLGPLLVLOAGPFLLRILITIP  
QSLDSWMTSLNPLGSGPVCIGONSPTSNSHSPSCPPICPGYRWMLRRFIIFIL  
LLCLFLVLLDYQMLPVCPLIPGSTTSTGPKCTCTTPAOGNSMFSPSCCKTKTDG  
NCTCPIFSSWAFKILMEWASVRFWSWLSLVFFVQWFWGLSFTVWLSAIMWMTWGP  
SLYSIVSPFILLPIFFCLMWYI"

## ORIGIN

Alignment Scores:  
Pred. No.: 7,24e-104 Length: 5882  
Score: 1746.00 Matches: 303  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-715-665-7\_COPY\_1\_303 (1-303) x AX154501 (1-5882)

QY 1 MetAspAlaMetLysArgGlyLeuCysCysValLeuLeuLeuCysGlyAlaValPheVal 20  
DB 1992 ATGGATGCAATGAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCAGTCTCGTT 2051  
QY 21 SerProSerAlaSerGluThrHisValThrGlyGlySerAlaGlyHisThrValSerGly 40  
DB 2052 TCGCCAGCGCTAGCGAACCACCGTCAACGGGGAGTGGCGGCCACACGTGTGTGGA 2111  
QY 41 PheValSerLeuLeuAlaProGlyAlaLysGlnAsnValGlnLeuLeuLeuThrGly 60  
DB 2112 TTGTGTAGCTCTCTCGACACAGCGCCCAAGCAGACGTCCAGCTGATCAACCAACGGC 2171  
QY 61 SerTrpHisLeuAsnSerThrAlaLeuAsnCysAsnAspSerLeuAsnThrGlyTrpLeu 80  
DB 2172 AGTTGGCACCCTCAATAGCAGCGCCCTGAATGCAATGATGCTCAACCGCTGTTG 2231  
QY 81 AlaGlyLeuPheTyrrHisHisPheAsnSerSerGlyCysProGluArgLeuAlaSer 100  
DB 2232 GCAGGGCTTTTCTATCACCACCAAGTTCAAATCTCTCAGGCTGCTCGAGAGGCTAGCCAGC 2291  
QY 101 CysArgProLeuThrAspPheAspGlnGlyTrpGlyProIleSerTyrrAlaAsnGlySer 120  
DB 2292 TCCGACCCCTTACCAGTTTGTACAGGGCTGGGGCCCTATCAGTTATGCCAACGGAGAC 2351  
QY 121 GlyProAspGlnArgProTyrrCysTrpHisTyrrProProLysPheProCysGlyIleValPro 140  
DB 2352 GGCCCCGACAGCGCCCTACTGCTGCACACTACCCCCCAAAACCTTGGCGTATTGTGCC 2411  
QY 141 AlaLysSerValCysGlyProValTyrrCysPheThrProSerProValValGlyThr 160  
DB 2412 GCGAAGAGTGTGTGTCGGGTATATTGTTTCACTCCAGCCCGCTGGTGGTGGGAACG 2471  
QY 161 ThrAspArgSerGlyAlaProThrTyrrSerTrpGlyGluAsnAspThrAspValPheVal 180  
DB 2472 ACCGACAGGTGGGGCGGCCCTACCTACAGCTGGGGTGAAAAATGATACGACGCTCTTCGTC 2531  
QY 181 LeuAsnAsnThrArgProProLeuGlyAsnTrpPheGlyCysThrTrpMetAsnSerThr 200  
DB 2532 CTYTAACAATACCAAGCCACCGCTGGGCAATTGGTTCCGTTGTACTTGGATGAACCTCACT 2591  
QY 201 GlyPheThrLysValCysGlyAlaProProCysValIleGlyAlaGlyAsnAsnThr 220  
DB 2592 GGATTCAACAAAGTGTGCGAGGCGCTCTCTGTGTATCGGAGGGCGGGCAACACACACC 2651  
QY 221 LeuHisCysProThrAspCysPheArgLysHisProAspAlaThrTyrrSerArgCysGly 240  
DB 2652 CTGACTGCCCCACTGATTGCTTCCGCAAGCATCCCGAGCGCCACATACTCTCGTGCGGC 2711  
QY 241 SerGlyProThrIleThrProArgCysLeuValAspTyrrProTyrrArgLeuThrHisTyrr 260  
DB 2712 TCCGGTCCCTGGATCACACCCAGGTGCTTGGTGCAGTACCCCGTATAGGCTTTTGGCATTAT 2771

QY 261 ProCysThrIleAsnTyrrThrIlePheLysIleArgMetTyrrValGlyGlyValGluHis 280  
DB 2772 CTTTGTACCATCACTACACCATATTTAAATCAGATGTACTGGGAGGGTTCGACAC 2831  
QY 281 ArgLeuGluAlaAlaCysAsnTrpThrArgGlyGluArgCysAspLeuGluAspArgAsp 300  
DB 2832 AGGCTGGAAGCTGCCTGCAACTGGAGCGGGCGAAACGTTGCGATCTCGAAGATAGGGAC 2891  
QY 301 ArgSerGlu 303  
DB 2892 AGGTCCGAG 2900  
RESULT 2  
AR004336  
LOCUS AR004336 1207 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 12 from patent US 5747241.  
ACCESSION AR004336  
VERSION AR004336.1 GI:3965215  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1207)  
AUTHORS Miyamura,T., Saito,I., Harada,S. and Honda,Y.  
TITLE Diagnostic reagent for hepatitis C  
JOURNAL Patent: US 5747241-A 12 05-MAY-1998;  
FEATURES Location/Qualifiers  
1..1207  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.9e-96 Length: 1207  
Score: 1620.50 Matches: 283  
Percent Similarity: 96.93% Conservative: 1  
Best Local Similarity: 96.59% Mismatches: 4  
Query Match: 92.81% Indels: 5  
DB: 6 Gaps: 1  
US-10-715-665-7\_COPY\_1\_303 (1-303) x AR004336 (1-1207)  
QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
DB 104 GTGCTGCTCTATTTCGCCGGCTC-----GACGGGAACCCACGTCACC 148  
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
DB 149 GGGGGAAGTGGCGGCACACTGTGTGATTTGTAGCTCTCTCGCACAGGCGCAAG 208  
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70  
DB 209 CAGAAGCTCAGCTATCAACCAACCGCAGTTGGCACCTCAATAGCAGCGCCCTGAAC 268  
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrrHisHisLysPheAsn 90  
DB 269 TGCATATGATAGCTCAACACCGCTGTTGGCAGGGCTTTTCTATCACCACAGTTCAAC 328  
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
DB 329 TCTTCAGGCTGTCTGAGAGGCTAGCAGCTGCCGACCCCTTACCGATTTTGACCAAGGC 388  
QY 111 TrpGlyProIleSerTyrrAlaAsnGlySerGlyProAspGlnArgProTyrrCysTrpHis 130  
DB 389 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACAGCGCCCTACTGCTGGCAC 448  
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrrCys 150  
DB 449 TACCCCCCAAAACCTTTCGGTATTGTGCCCGCGAAGAGTGTGTGCTCGGTATATTGC 508  
QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrrSer 170  
DB 509 TTCACCTCCAGCCCCCGTGGTGGTGGGACGACGACAGGTTCGGCGCGGCCCTACCTACAGC 568

```
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 569 TGGGGTGAATAATGATACGACGCTCTTGTCTCTTAACAATACACGGCCACCGCTGGGCAAT 628
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrIysValCysGlyAlaProPro 210
Db 629 TGGTTCGGTGTACTCGATGAGAACTCAACTGGATTCAACAAAGTGTGCGGAGCGCTCCT 688
Qy 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 689 TGTGTATCGAGAGGGGGGGGCAACACACCTTGCACTGCCCACTGATTGCTTCGGCAAG 748
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 749 CATCCGAGCGCACATACTCTCGTGGCTCCGCTCCCTCGATCACACCCAGGTGCTG 808
Qy 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 809 GTCGACTACCGGTATAGGCTTTGGCATTATCTCTTGTACCACTACACCATATTTAAA 868
Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaCysAsnTrpThrArg 290
Db 869 ATCAGGATGTACGTGGAGGGGTGCAACACAGGCTGGAAGTCTGCAACTGGACCGG 928
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 929 GCGGAACGTTGCGATCTGGAAGACAGGACAGGTCGCGAG 967

RESULT 3
LOCUS AR006827 1207 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 12 from patent US 5750331.
ACCESSION AR006827
VERSION AR006827.1 GI:3966311
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1207)
AUTHORS Miyamura, T., Saito, I., Harada, S. and Honda, Y.
TITLE Diagnostic reagent for hepatitis C
JOURNAL Patent: US 5750331-A 12 12-MAY-1998;
FEATURES
    source
    1. .1207
    /organism="unknown"
    /mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1 9e-96 Length: 1207
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 6 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x AR006827 (1-1207)
Qy 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 104 GTGCTGTGCTGTTATTTGCGCGGCTC-----GACGCGGAACCCACGTCAAC 148
Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 149 GGGGGAAGTCCGGCCCACTGTGTCTGGATTGTTAGCTCTCTCGCACCGGCGCAAG 208
Qy 51 GluAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 209 CAGAACGTCAGCTGATCAACACACAGCGAGTGGCACCTCAATAGCAGCGCTGAC 268
Qy 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
```

```
Db 269 TGCATATGATAGCTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCAAGTTCAC 328
Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 329 TCITCAGGCTGTCTGAGAGGTAGCCAGCTGCCGACCCCTTACCGATTTTTCACAGGCG 388
Qy 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 389 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACAGCGCCCTACTGTCTGGCAC 448
Qy 131 TyrProIleCysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 449 TACCCCCCAAAACCTTCGGGTATTGTCCCGCAAGAGTGTGTGGTCCGCTATATTTCG 508
Qy 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 509 TTCCTCCAGCCCCGCTGGTGGGAACGACCGAGGTGGGGCGGCCCACTTACAGC 568
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 569 TGGGGTGAAATAATGATACGACGCTCTTGTCTCTTAACAATACACGGCCACCGCTGGCAAT 628
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrIysValCysGlyAlaProPro 210
Db 629 TGGTTCGGTGTACTCGATGAGAACTCAACTGGATTCAACAAAGTGTGCGGAGCGCTCCT 688
Qy 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 689 TGTGTATCGAGAGGGGGGGGCAACACACCTTGCACTGCCCACTGATTGCTTCGGCAAG 748
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 749 CATCCGAGCGCACATACTCTCGTGGCTCCGCTCCCTCGATCACACCCAGGTGCTG 808
Qy 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 809 GTCGACTACCGGTATAGGCTTTGGCATTATCTCTTGTACCACTACACCATATTTAAA 868
Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaCysAsnTrpThrArg 290
Db 869 ATCAGGATGTACGTGGAGGGGTGCAACACAGGCTGGAAGTCTGCAACTGGACCGG 928
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 929 GCGGAACGTTGCGATCTGGAAGACAGGACAGGTCGCGAG 967

RESULT 4
LOCUS AR031233 1210 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 35 from patent US 5866139.
ACCESSION AR031233
VERSION AR031233.1 GI:5945522
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1210)
AUTHORS Brechot, C., Krams, D., and Porchon, C.
TITLE Nucleotide and peptide sequences of a hepatitis C virus isolate, diagnostic and therapeutic applications
JOURNAL Patent: US 5866139-A 35 02-FEB-1999;
FEATURES
    source
    1. .1210
    /organism="unknown"
    /mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1 91e-96 Length: 1210
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
```

DB:	6	Gaps:	1
US-10-715-665-7_COPY_1_303 (1-303) x AR031233 (1-1210)			
QY	11	ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr	30
DB	122	GTGCTGCTGCTATTTCGGCGGCTC-----GACGCGGAAACCCACGTCACC	166
QY	31	GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys	50
DB	167	GGGGGAAGTGGCGGCACACTGTCTGGATTGTGTAGCTCTCGCACCGCGCCAAAG	226
QY	51	GlnAenValGlnLeuLeuAenThrAsnGlySerTrpHisLeuAenSerThrAlaLeuAen	70
DB	227	CAGAAGCTCAGCTATCAACACCGAGTGGCACTCAATAGCACGGCTCTGAAC	286
QY	71	CysAenAspSerLeuAenThrGlyTrpLeuAlaGlyLeuPheThrHisLysPheAen	90
DB	287	TGCNATGATAGCTTAACACCGGCTGTTGGCAGGCTTTCTATCACCACAAGTTCAAC	346
QY	91	SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheaspGlnGly	110
DB	347	TCTTCAGGCTGCTCAGAGGCTAGCCAGCTGCGACCCCTTACCGATTTCACGAGGC	406
QY	111	TrpGlyProIleSerTrpAlaAenGlySerGlyProAspGlnArgProTrpHis	130
DB	407	TGGGGCCCTATCATGTTATGCCAACGGAAGCGCCCGACCGGCTCTACTGCTGCAC	466
QY	131	TyrProProLysProCysGlyLeuValProAlaLysSerValCysGlyProValTyrCys	150
DB	467	TGCGTATCGGAGGGCGGCAACACACCTGCACTGCCCTGCTGCTGCTGCTGCTGCTGCT	526
QY	151	HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu	170
DB	527	CATCCGGACGCCACATACCTCTCGGTGGGCTCCGGTCCCTGGATCACCCAGTGCCTG	586
QY	171	ValAspTyrProTrpArgLeuTrpHisTyrProCysThrIleAenThrIlePheLys	190
DB	587	GTGCTACCCGATAGGCTTTGGCATTTATCCTTTGTCATCACTCACTACCATATTATA	646
QY	271	IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAenThrTrpArg	290
DB	887	ATCAGGATGTACGTGGAGGGTGGAAACACAGGCTGGAAAGCTGCCGCACTGGACGCG	946
QY	291	GlyGluArgCysAspLeuGluAspArgAspArgSerGlu	303
DB	947	GCGGAACGTTGCGATCTGGAAGACAGGGACAGGTCCGAG	985
RESULT 5			
LOCUS	AR145049	1210 bp	DNA linear PAT 08-AUG-2001
DEFINITION	Sequence 35 from patent US 6210962.		
ACCESSION	AR145049		
VERSION	AR145049.1 GI:15106916		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		

REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1210)		
TITLE	Brechot,C., Kremsdorff,D. and Porchon,C. Nucleotide and peptide sequences of an isolate of the hepatitis C virus, diagnostic and therapeutic applications thereof		
JOURNAL	Patent: US 6210962-A 35 03-APR-2001;		
FEATURES	Location/Qualifiers		
source	1..1210		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Alignment Scores:	1..91e-96	Length:	1210
Pred. No.:	1620.50	Matches:	283
Score:	96.93%	Conservative:	1
Percent Similarity:	96.59%	Mismatches:	4
Best Local Similarity:	92.81%	Indels:	5
Query Match:	6	Gaps:	1
DB:			
US-10-715-665-7_COPY_1_303 (1-303) x AR145049 (1-1210)			
QY	11	ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr	30
DB	122	GTGCTGCTGCTATTTCGGCGGCTC-----GACGCGGAAACCCACGTCACC	166
QY	31	GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys	50
DB	167	GGGGGAAGTGGCGGCACACTGTCTGGATTGTGTAGCTCTCGCACCGCGCCAAAG	226
QY	51	GlnAenValGlnLeuLeuAenThrAsnGlySerTrpHisLeuAenSerThrAlaLeuAen	70
DB	227	CAGAAGCTCAGCTATCAACACCGAGTGGCACTCAATAGCACGGCTCTGAAC	286
QY	71	CysAenAspSerLeuAenThrGlyTrpLeuAlaGlyLeuPheThrHisLysPheAen	90
DB	287	TGCNATGATAGCTTAACACCGGCTGTTGGCAGGCTTTCTATCACCACAAGTTCAAC	346
QY	91	SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheaspGlnGly	110
DB	347	TCTTCAGGCTGCTCAGAGGCTAGCCAGCTGCGACCCCTTACCGATTTCACGAGGC	406
QY	111	TrpGlyProIleSerTrpAlaAenGlySerGlyProAspGlnArgProTrpHis	130
DB	407	TGGGGCCCTATCATGTTATGCCAACGGAAGCGCCCGACCGGCTCTACTGCTGCAC	466
QY	131	TyrProProLysProCysGlyLeuValProAlaLysSerValCysGlyProValTyrCys	150
DB	467	TGCGTATCGGAGGGCGGCAACACACCTGCACTGCCCTGCTGCTGCTGCTGCTGCTGCT	526
QY	151	HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu	170
DB	527	CATCCGGACGCCACATACCTCTCGGTGGGCTCCGGTCCCTGGATCACCCAGTGCCTG	586
QY	171	ValAspTyrProTrpArgLeuTrpHisTyrProCysThrIleAenThrIlePheLys	190
DB	587	GTGCTACCCGATAGGCTTTGGCATTTATCCTTTGTCATCACTCACTACCATATTATA	646
QY	271	IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAenThrTrpArg	290
DB	887	ATCAGGATGTACGTGGAGGGTGGAAACACAGGCTGGAAAGCTGCCGCACTGGACGCG	946
QY	291	GlyGluArgCysAspLeuGluAspArgAspArgSerGlu	303
DB	947	GCGGAACGTTGCGATCTGGAAGACAGGGACAGGTCCGAG	985



Db	496	TGGGGTGAATAATGATACGGACGCTCTTCGTCTCTTAAACAATACACAGGCCACCGCTGGGCAAT	555
Qy	191	TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro	210
Db	556	TGGTTCGGTTGTACTCGTGAATGAACCTCAACTGGATTCCACAAAGTGTGGCGAGCGCTCTCT	615
Qy	211	CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys	230
Db	616	TGTGTCAATCGAGGGGGGGGCAACAACCTTGCACTGCCCACTGGATTGCTTCGCGAAG	675

Db 616 TGTGTCATCGAGGGGGGCAACAACCCCTGCCTGCCACTGCTTGTCTTCGCAAG 675

Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250

**Qy**      231 HisProaspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
**Db**      676 CATCCGACGCCAATACTCTCGGTGGGCCTCCGGTCCCTGGATCACCCCAGGTGCCTTG 735

DB	6/8	CAATCCGACGCCACATATCTCTTCGGTTCGGGCTCCCTGGATCACACCCAGGTGCTTG	270
Qy	251	ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys	270
DB	7/8	GTCCGATTAACCGGTATAGCTTTGGCATTTATCTCTGTACCAATCAACTACA	795
DB	7/8	GTCCGATTAACCGGTATAGCTTTGGCATTTATCTCTGTACCAATCAACTACA	795

Db	736	GTCGACTACCGTAGCTTTGGCATTTATCCTTGTACCATCAACTACATATATTTAA	795
Qy	271	IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg	290

Accession	Protein	Accession	Protein
Qy	IleargmettyrvalgalvglyvvalgluhisargleugluAlaAlaCysasntprthrarg	290	299
Db	796 ATCAGSAGTGACGTGGGAGGGGTCGAGCACAGGCTGGAGCTGCCTGCAACTGGACGGC	855	

**Qy** 291 GlyClnArgCysAspLeuGluAspArgAspArgSerGlu 303  
**Dd** 856 GGCAGACGTTCGGATCTGGAGATAGGGACAGGTCGAG 894	

Db 856 GGCGAACGTTGCGATCTGGAGATAGGACAGGTCGGAG 894

RESULT 7  
BD132859

BD132859	BD132859	1989 bp	DNA	linear	PAT 18-SEP-2000
LOCUS					
DEFINITION	Intracellular production of hepatitis C E1 and E2 truncated polypeptides.				

DEFINITION Intracellular production of hepatitis C, E1 and E2 circulates polypeptides.

ACCESSION BD132859

VERSION BD132859.1 GI:23227804

KEYWORDS JP 2002504810-A/2.

VERSION BUI32859.1  
 KEYWORDS JP 202504810-A/2.  
 SOURCE unidentified  
 ORGANISM unidentified  
 unclassified

ORGANISM  
unclassified  
unclassified  
1 (bases 1 to 1989)  
Houghton, M., Choo, Q. L., Abrignani, S., Chien, D., Selby, M. and  
Glazer, F.

**AUTHORS** Houghton, M., Choo, Q.L., Abrignani, S., Chien, D., Seiby, M. and Glazer, E.  
**TITLE** Intracellular production of hepatitis C E1 and E2 truncated  
**JOURNAL** Patent: JP 2002504810-A 2 12-FEB-2002;  
CUTPON COOP

JOURNAL	COMMENT
Patent: JP 2002504810-A 2 12-FEB-2002;	
CHIRON CORP	OS Hepatitis virus (hepatitis C virus)
PN JP 2002504810-A/2	
50 12 FEB 2002	

PN	JP 2002504810-A/2
PD	12-FEB-2002
PF	06-MAY-1998 JP 1998548360
PR	06-MAY-1997 US 60/045675

PR 06-MAY-1997 US 60/045675  
PI MICHAEL Houghton, QUI LIM CHOO, SERGIO ABRIGNANI, DAVID CHIEN, MARI  
PI SELBY,

PI SELBY,  
PI EDWARD GLAZER  
PC C12N15/40,C07K14/18,A61K39/29,G01N33/50  
CC

FEATURES	Location/Qualifiers
CC	Location/Qualifiers
FFH	(1). . (1989).
FT	Location/Qualifiers

FEATURES
source
1..1989
Location/Qualifiers
17..(1989).
CDS
f1
/organism="unidentified"
/mol type="genomic DNA"

```

ORIGIN
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

```

ORIGIN	
Alignment Scores:	
Pred. No.:	3,23e-96
Score:	1620.50
Length:	1989
Matches:	283

Pred. No.:	3.33e-96	Length:	1989
Score:	1620.50	Matches:	283
Percent Similarity:	96.93%	Conservative:	1
Best Local Similarity:	96.59%	Mismatches:	4
Cons. Match:	92.91%	Indels:	5

Best Local Similarity:	96.59%	Mismatches:	4
Query Match:	92.81%	Indels:	5
DB:	6	Gaps:	1

US-10-715-665-7\_COPY\_1\_303 (1-303) x BD132859 (1-1989)



Db	1303	TGGTTCGGTGTGTAACCTCAACTGGGATTCACCAAGTGTGCGAGCGCCTCCT	1362			
Qy	211	CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys	230			
Db	1363	TGTGTCAATCGAGGGGGGCGCAACACACCTGCAGCTGCCCACTGATTGCTTCGCGAAG	1422			
Qy	231	HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu	250			
Db	1423	CATCCGAGCCACATACCTCTCGGTGGGTCTCGGTCTCCGTGGATCACACCGAGTGCCTG	1482			
Qy	251	ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys	270			
Db	1483	GTCAGCTACCCGTATAGGCTTTGGCATTAATCCCTGTACCATCAACATACCATATTATAA	1542			
Qy	271	IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg	290			
Db	1543	ATCAGGATGTACGTGGAGGGGTGGAACACAGCTCGAAGCTGCCTGCACTGGACGCGG	1602			
Qy	291	GlyGluArgCysAspLeuGluAspArgAspArgSerGlu	303			
Db	1603	GGCGAAGCTTGCATCTGGAAGACAGGACAGGTCCGAG	1641			
RESULT 10						
AR118728						
LOCUS	AR118728	8987 bp	DNA linear PAT 16-MAY-2001			
DEFINITION	Sequence 137 from patent US 6150087.					
ACCESSION	AR118728					
VERSION	AR118728.1 GI:14100638					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 8987)					
AUTHORS	Chien,D.Y.					
TITLE	NANBV diagnostics and vaccines					
JOURNAL	Patent: US 6150087-A 137 21-NOV-2000;					
FEATURES	Location/Qualifiers					
source	1..8987					
	/organism="unknown"					
	/mol_type="unassigned DNA"					
ORIGIN						
Alignment Scores:						
Pred. No.:	1.6e-95	Length:	8987			
Score:	1620.50	Matches:	283			
Percent Similarity:	96.93%	Conservative:	1			
Best Local Similarity:	96.59%	Mismatches:	4			
Query Match:	92.81%	Indels:	5			
DB:	6	Gaps:	1			
US-10-715-665-7_COPY_1_303 (1-303) x AR118728 (1-8987)						
Qy	11	ValIleuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr	30			
Db	1120	GTGCTGCTGCTATTTCGGCGGCTC-----GACGCGGAACCCACGTCACC	1164			
Qy	31	GlyGlySerAlaGlyHisThrValSerGlyPheValSerIleuLeuAlaProGlyAlaLys	50			
Db	1165	GGGGGAAGTCCCGCCACACTGTGTCTGGATTGTTAGCTCTCTCGCACCGGCGCAAG	1224			
Qy	51	GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisIleAsnSerThrAlaLeuAsn	70			
Db	1225	CAGAACGTCAAGCTGATCAACACCAACGCGAGTTGGCACTCAATAGCAGGCGCCCTGAAC	1284			
Qy	71	CysAsnAspSerIleuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisIleLysPheAsn	90			
Db	1285	TGCAATGATAGCTCAACACCGCTGTTGGCAGGCGCTTTCTATCACCAAGTTCAC	1344			
Qy	91	SerSerGlyCysProGluArgIleuAlaSerCysArgProIleuThrAspPheAspGlnGly	110			
Db	1345	TCATTACGGTGTCTCAGAGGCTAGCAGCTCGGACCCCTACCGATTTGACCAAGGC	1404			

QY 111 TtpGlyProIleSerTyrAlaIleGlySerGlyProAspGlnArgProTyrCysTrpHis 130  
 DB 1405 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACAGCGCCCTACTGCTGGCAC 1464  
 QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
 DB 1465 TACCCCCCAAACTTGGCGTATTTGGCCCGGAAGAGTGTGTGGTCCGGTATATTGC 1524  
 QY 151 PheThrProSerProValValGlyThrAspArgSerGlyAlaProThrTyrSer 170  
 DB 1525 TTCACTCCCGAGCCCGTGGTGGGACAGCAGCTCGGCGCGCCACCTACAGC 1584  
 QY 171 TtpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
 DB 1585 TGGGGTGAATATGATACGACGCTTTCGTCCTTAACATACAGGCCACCGCTGGCAAT 1644  
 QY 191 TtpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
 DB 1645 TGGTTCCGTTGTACCTGGATGAACCACTGGATTACCAAAAGTGTGGGAGCGCTCCT 1704  
 QY 211 CysValIleGlyCysGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 DB 1705 TGTGTATCGAGGCGCGGCAACACACCTGCACTGCCCTGCACTGCCCTGCACTGCCCTG 1764  
 QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTyrTrpIleThrProArgCysLeu 250  
 DB 1765 CATCCGAGCGCACATACCTCTCGTGGCGCTCCGGTCCCTGGATCACCCAGGTGCCCTG 1824  
 QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
 DB 1825 GTCCAGTACCCGTATAGGCTTTGGCATTTATCTTGTACCATCAACTACACCATATTAAA 1884  
 QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
 DB 1885 ATCAGGATGTAGTGGGAGGGTGGAAACACAGCTGGAAGCTGCCCTGCAACTGGACGCG 1944  
 QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
 DB 1945 GCGGAACGTTGGATCTGGAAGACAGGACAGGTCCGAG 1983

RESULT 11  
 AR118722 LOCUS AR118722 9185 bp DNA linear PAT 16-MAY-2001  
 DEFINITION Sequence 122 from patent US 6150087.  
 ACCESSION AR118722  
 VERSION AR118722.1 GI:14100632  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 9185)  
 AUTHORS Chien,D.Y.  
 TITLE NANBV diagnostics and vaccines  
 JOURNAL Patent: US 6150087-A 122 21-NOV-2000;  
 FEATURES Location/Qualifiers  
 source 1..9185  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.64e-95 Length: 9185  
 Score: 1620.50 Matches: 283  
 Percent Similarity: 96.93% Conservative: 1  
 Best Local Similarity: 96.59% Mismatches: 4  
 Query Match: 92.81% Indels: 5  
 DB: 6 Gaps: 1  
 US-10-715-665-7\_COPY\_1\_303 (1-303) x AR118722 (1-9185)  
 QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
 DB 1439 GTGCTGCTGCTATTTCGCGCGTC-----GACGCGAAACCCACGTCACC 1483

QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
 DB 1484 GGGGGAAGTGGCGGCCACACTGTGTGATTTGTAGCTCTCGACACGAGCCCGAAG 1543  
 QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70  
 DB 1544 CAGAACGTCCAGCTGATCAACACCAACGCGAGTTGGCACCTCAATAGCACGSCCTGAAC 1603  
 QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90  
 DB 1604 TGCAATGATAGCTCAACACCGCTGTGGCAGGCTTTCTCTATCACCCACAAGTTCAC 1663  
 QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
 DB 1664 TCTTCAGGCTGTCTGAGAGGCTAGCAGCTGCCACCCCTTACCGATTTTGCACGAGGC 1723  
 QY 111 TtpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130  
 DB 1724 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACAGCGCCCTACTGCTGGCAC 1783  
 QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
 DB 1784 TACCCCCCAAACTTGGCGTATTTGGCCCGCAAGAGTGTGTGGTCCGGTATATTGC 1843  
 QY 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170  
 DB 1844 TTCACCTCCAGCGCCGCTGTGGTGGAAACGACACAGCTGGGCGCGCCACCTACAGC 1903  
 QY 171 TtpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
 DB 1904 TGGGGTGAATATGATACGACGCTTTCGTCCTTAAACAATACACAGGCCACCGCTGGCAAT 1963  
 QY 191 TtpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
 DB 1964 TGGTTCGGTGTGTACCTGGATCAACTCAACTGGATTCACCAAAAGTGTGGGAGCGCTCCT 2023  
 QY 211 CysValIleGlyGlyValGlyValGluHisCysProThrAspCysPheArgLys 230  
 DB 2024 TGTGTATCGGAGGGCGGCGCAACACACCTGCACTGCCCGCTGATTTGCTTCGCGAAG 2083  
 QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTyrTrpIleThrProArgCysLeu 250  
 DB 2084 CATCCGAGCGCCACATACCTCTCGGTGGCTCCGCTCCCTGGATCACACCGAGTGCCTG 2143  
 QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
 DB 2144 GTCCAGTACCCGCTATAGGCTTTGGCATTTATCTTGTACCATCAACTACACCATATTAAA 2203  
 QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
 DB 2204 ATCAGGATGTAGTGGGAGGGTGGAAACACAGCTGGAAGCTGCTGCAACTGGACGCGG 2263  
 QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
 DB 2264 GCGGAACGTTGGATCTGGAAGACAGGACAGGTCCGAG 2302  
 RESULT 12  
 AR118723/c LOCUS AR118723 9185 bp DNA linear PAT 16-MAY-2001  
 DEFINITION Sequence 123 from patent US 6150087.  
 ACCESSION AR118723  
 VERSION AR118723.1 GI:14100633  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 9185)  
 AUTHORS Chien,D.Y.  
 TITLE NANBV diagnostics and vaccines  
 JOURNAL Patent: US 6150087-A 123 21-NOV-2000;  
 FEATURES Location/Qualifiers  
 source 1..9185



```
Db      2024  TGTGTTCATCGAGGGGGCGGCAACACCCCTGCACCTGCCACCTGATTGCTTCGCGAAG 2083
Qy      231  HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db      2084  CATCGGACGCCACATACTCTCGGTGGGCTCCGCTCCCTGGATCACACCCAGGTGCCTG 2143
Qy      251  ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db      2144  GTCGACTACCCGCTATAGGCTTTGGCATATATCTTGTTACCATCAACTACACCATATTAAA 2203
Qy      271  IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
Db      2204  ATCAGGATGTAGTGGAGGGGTGCAACACAGCTGGAAGCTGCTGCAACTGGACGGCG 2263
Qy      291  GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db      2264  GCGGAACGTTGCGATCTGGAACACAGGACAGGTCGCGAG 2302

RESULT 14
LOCUS      BD091382                      9185 bp    DNA        linear    PAT 27-AUG-2002
DEFINITION HCV cultivation method in eucaryotic cells.
ACCESSION  BD091382
VERSION     BD091382.1 GI:22636993
KEYWORDS   JP 2001314192-A/3.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 9185)
AUTHORS    Weiner A.J., Steimer K.S. and Houghton M.
TITLE      HCV cultivation method in eucaryotic cells
JOURNAL    Patent: JP 2001314192-A 3 13-NOV-2001;
           CHIRON CORP
COMMENT    OS Homo sapiens (human)
           PN JP 2001314192-A/3
           PD 13-NOV-2001
           PF 15-MAR-2001 JP 2001075114
           PR 25-AUG-1989 US 398667
           PI AMY J WEINER, KATHLEEN S STEIMER, MICHAEL HOUGHTON PC
           C12N15/09, C12N5/10, C12N7/00// (C12N7/00, C12R1:93), C12N15/00, PC
           C12N5/00
           CC HCV cultivation method in eucaryotic cells
           FH Key
           FT source
           FT Location/Qualifiers
           1..9185
             /organism="Homo sapiens (human)"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"

FEATURES
   source
       1..9185
       Location/Qualifiers
           1..9185
           Length: 9185
           Pred. No.: 1
           Score: 1620.50
           Percent Similarity: 96.93%
           Best Local Similarity: 96.59%
           Query Match: 92.81%
           DB: 6
           Matches: 283
           Conservative: 1
           Indels: 5
           Gaps: 1

ORIGIN
Alignment Scores:
Pred. No.: 1
Score: 1620.50
Percent Similarity: 96.93%
Best Local Similarity: 96.59%
Query Match: 92.81%
DB: 6

US-10-715-665-7_COPY_1_303 (1-303) x BD091382 (1-9185)
Qy      11  ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db      1439  GTGCTGCTGCTATTTCGCGCGTC-----GACGCGAAACCCACGCTCACC 1483
Qy      31  GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db      1484  GGGGGAAGTCCGCGCCACACTGTGTCTGGATTGTTAGCTCTCTCCACACAGCGCGCAAG 1543
Qy      51  GlnAsnValGlnLeuIleAsnThrAenGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db      1544  CAGAAGCTCCAGCTGATCAACACCAACCGAGTTGGCACCTCAATAGCACGCGCCCTGAAC 1603
Qy      71  CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
Db      1604  TGCATATGATAGCCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCAAGTTCAAC 1663
Qy      91  SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db      1664  TCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTATTGACGAGGC 1723
Qy      111  TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db      1724  TGGGCGCCTATCAGTTATGCCAAGCGAGCGCGCCCGACAGCGCCCTACTGCTGGCAC 1783
Qy      131  TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db      1784  TACCCCCCAAAACCTTGGCGTATTGTGCCCGGAAGAGTGTGTGGTCCGGTATATTGC 1843
Qy      151  PheThrProSerProValValGlyThrThrAspArgSerGlyValAlaProThrTyrSer 170
Db      1844  TTCACTCCAGCCCGCTGGTGGGAACGACGACAGGTCCGGCGCGCCACCTACAGC 1903
Qy      171  TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db      1904  TGGGCTGAAATGATACGAGCTCTTCGCTTAAACAATACCAAGCCACCCGCTGGGCAAT 1963
Qy      191  TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db      1964  TGGTTCGGTGTGTACCTGGATGAACTCAACTGGATTACCAAGTGTGGAGCGGCTCTCT 2023
Qy      211  CysValIleGlyGlyValAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db      2024  TGTGTATCGAGGGGGCGGCAACCAACCTGCACCTGCCACCTGATTCCTCCGCAAG 2083
Qy      231  HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db      2084  CATCGGACGCCACATACTCTCGGTGGGCTCCGCTCCCTGGATCACACCCAGGTGCCTG 2143
Qy      251  ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db      2144  GTCGACTACCCGCTATAGGCTTTGGCATATATCTTGTTACCATCAACTACACCATATTAAA 2203
Qy      271  IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
Db      2204  ATCAGGATGTAGTGGAGGGGTGCAACACAGCTGGAAGCTGCTGCAACTGGACGGCG 2263
Qy      291  GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db      2264  GCGGAACGTTGCGATCTGGAACACAGGACAGGTCGCGAG 2302

RESULT 15
LOCUS      AR166930                      9379 bp    DNA        linear    PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6284249.
ACCESSION  AR166930
VERSION     AR166930.1 GI:16243325
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 9379)
AUTHORS    Barban V.
TITLE      Fusion polypeptide having the C protein and E1 protein of hepatitis
JOURNAL    Patent: US 6284249-A 1 04-SEP-2001;
           Location/Qualifiers
           1..9379
           /organism="unknown"
           /mol_type="unassigned DNA"
           source
ORIGIN
Alignment Scores:
Pred. No.: 1
Score: 168e-95
Length: 9379
```



Score: 1620.50 Matches: 283  
 Percent Similarity: 96.93% Conservative: 1  
 Best Local Similarity: 96.59% Mismatches: 4  
 Query Match: 92.81% Indels: 5  
 DB: 6 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x AR166930 (1-9379)

Qy	11	ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr	30
Db	1439	GTGCTGCTGCTATTGGCGGCTC-----GACGGGAACCCACGTCACC	1483
Qy	31	GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys	50
Db	1484	GGGGGAAGTCCCGCCACACACTGTGCTGGATTGTAGCTCTCGCACCGGCGCAAG	1543
Qy	51	GlnAsnValGlnLeuIleAsnThrAsnGlySerThrHisLeuAsnSerThrAlaLeuAsn	70
Db	1544	CAGAAGCTCCAGCTGATCAACACACCGGAGTGGCACCTCAATAGCACGGGCCCTGAAC	1603
Qy	71	CysAsnAspSerLeuAsnThrGlyTyrLeuAlaGlyLeuPheTyrHisHisLysPheAsn	90
Db	1604	TGCAATGATAGCTTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCAAGTTCAAC	1663
Qy	91	SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly	110
Db	1664	TCCTCAGGCTGTCTCAGAGGCTAGCCAGCTGCCAGCCCTTACCGATTTTGACGAGGC	1723
Qy	111	TyrGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTyrHis	130
Db	1724	TGGGGCCCTATCAGTTATGCTCAACGGAAGCGGCCCGCCAGCGCCCTACTGCTGGCAC	1783
Qy	131	TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys	150
Db	1784	TACCCCCCAAACTTGGCGTATTGTGCCCGCGAAGAGTGTGTGGTCCGGTATATTGC	1843
Qy	151	PheThrProSerProValValValGlyThrAspArgSerGlyAlaProThrTyrSer	170
Db	1844	TTCACTCCCAAGCCCGTGGTGGGAAACGACGACAGGTCCGGCGCGCCACCTACAGC	1903
Qy	171	TyrGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn	190
Db	1904	TGGGGTGAAATGATACGGACGTCTTCGTCTTAAACAATACCGAGCCCGCTGGGCAAT	1963
Qy	191	TyrPheGlyCysThrTyrMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro	210
Db	1964	TGGTTCGGTTGTACCTGGATGAACCTCACTGGATTCAACCAAGTGTGGGAGCGCTCCT	2023
Qy	211	CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys	230
Db	2024	TGTGTCACTCGAGGGGCGGCAACACACCCCTGCACCTGCCCCACTGATTGCTTCCGCAAG	2083
Qy	231	HisProAspAlaThrTyrSerArgCysGlySerGlyProThrIleThrProArgCysLeu	250
Db	2084	CATCCGGACCCACATACCTCGGTGGGCTCCGCTCCCTGGATCAACCCAGGTGCTG	2143
Qy	251	ValAspTyrProTyrArgLeuTyrHisTyrProCysThrIleAsnTyrThrIlePheLys	270
Db	2144	GTGCACTACCCGATAGGCTTTGGCATATCTCTTGTACCATCACTACCATATTATAA	2203
Qy	271	IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnThrThrArg	290
Db	2204	ATCAGGATGTACGTGGAGGGGTGGAACACAGGCTGGAGCTGCCTGCAACTGGACGCGG	2263
Qy	291	GlyGluArgCysAspLeuGluAspArgAspArgSerGlu	303
Db	2264	GGCGAACGTTGCGATCTGGAAGACAGGGACAGGTCCGAG	2302

Search completed: April 15, 2005, 08:04:41  
 Job time : 4932 secs

This page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 15, 2005, 02:41:03 ; Search time 611 Seconds  
(without alignments)  
2935.649 Million cell updates/sec

Title: US-10-715-665-7\_COPY\_1\_303

Perfect score: 1746

Sequence: 1 MDAMKGLCCVLLCGAVFV.....AACNWTGRGCDLDRDRSE 303

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10715665/runat\_13042005\_155821\_28129/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_16Dec04 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=200 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US10715665 @CGN 1 1 470 @runat\_13042005\_155821\_28129 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq\_16Dec04:\*  
1: Geneseq\_1980s:\*  
2: Geneseq\_1980s:\*  
3: Geneseq\_2000s:\*  
4: Geneseq\_2001as:\*  
5: Geneseq\_2001bs:\*  
6: Geneseq\_2002as:\*  
7: Geneseq\_2002bs:\*  
8: Geneseq\_2003as:\*  
9: Geneseq\_2003bs:\*  
10: Geneseq\_2003cs:\*  
11: Geneseq\_2003ds:\*  
12: Geneseq\_2004as:\*  
13: Geneseq\_2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Resid	Score	Query Match	Length	DB ID	Description
1746	100.0	5882	5	AAD06793	Aad06793 Plasmid p
20.5	92.8	1207	2	AAQ40330	AAQ40330 Sequence
7.5	92.8	1914	8	ABZ70946	ABZ70946 Hepatitis
5	92.8	1914	8	ACC71824	ACC71824 HCV-1 E1/
5	92.8	1914	12	ADO00769	ADO00769 HCV-1 E1/

6	1620.5	92.8	1989	2	AAT18246	Aat18246 HCV1 E2/N
7	1620.5	92.8	1989	2	AAV81371	Av81371 Hepatitis
8	1620.5	92.8	8316	2	AAQ05955	Aa05955 Hepatitis
9	1620.5	92.8	8316	3	AAV52396	Aa52396 CDNA sequ
10	1620.5	92.8	8316	12	ADN35977	Adn35977 HCV CDNA
11	1620.5	92.8	9185	2	AAQ05956	Aa05956 Sense str
12	1620.5	92.8	9185	2	AAQ10566	Aa10566 Hepatitis
13	1620.5	92.8	9185	2	AAQ00459	Aa00459 Hepatitis
14	1620.5	92.8	9185	2	AAQ26737	Aa26737 Nucleotid
15	1620.5	92.8	9185	2	ADP66068	Adp66068 Hepatitis
16	1620.5	92.8	9185	3	AAV5297	Aa5297 Sense str
17	1620.5	92.8	9400	2	AAQ21744	Aa21744 Compiled
18	1620.5	92.8	9401	2	AAQ12710	Aa12710 Hepatitis
19	1620.5	92.8	9401	2	AAV99981	Av99981 HCV poly
20	1620.5	92.8	9401	2	AAV09989	Av09989 HCV poly
21	1620.5	92.8	9401	6	AAQ35043	Aa35043 Hepatitis
22	1620.5	92.8	9401	12	ADL23106	Adl23106 Hepatitis
23	1620.5	92.8	9401	13	ADR29358	Adr29358 Hepatitis
24	1617.5	92.6	9185	12	ADN35979	Adn35979 HCV CDNA
25	1612.5	92.4	9379	2	AAQ36209	Aa36209 Composite
26	1608.5	92.1	9618	11	ADN33102	Adn33102 Hepatitis
27	1580.5	90.5	9133	2	AAZ07656	Aaz07656 Nucleotid
28	1530.5	87.7	7178	12	ADL72979	Adl72979 Hepatitis
29	1526	87.4	5125	2	AAQ47196	Aa47196 Plasmid p
30	1526	87.4	5323	2	AAQ47195	Aa47195 Plasmid p
31	1523.5	87.3	1207	2	AAQ40331	Aa40331 Sequence
32	1521.5	87.1	7551	12	ADL72981	Adl72981 Hepatitis
33	1521.5	87.1	7754	12	ADL72977	Adl72977 Hepatitis
34	1521.5	87.1	9401	2	AAQ41882	Aa41882 Hepatitis
35	1521.5	87.1	9416	2	AAV59378	Av59378 Hepatitis
36	1521.5	87.1	9416	6	ABK87300	Abk87300 CDNA enco
37	1521.5	87.1	9416	8	ACA62483	Ac62483 HCV-H CDN
38	1521.5	87.1	9622	10	AAV54424	Aa54424 Hepatitis
39	1518.5	87.0	9365	6	AAQ25518	Aa25518 Hepatitis
40	1515.5	86.8	9646	2	AAV59361	Av59361 Hepatitis
41	1515.5	86.8	9646	6	ABK87285	Abk87285 CDNA enco
42	1515.5	86.8	9646	8	ACA62466	Ac62466 HCV H77 c
43	1515.5	86.8	11674	3	AAQ36210	Aa36210 Nucleotid
44	1515.5	86.8	12980	2	AAV59364	Av59364 Hepatitis
45	1515.5	86.8	12980	6	ABK87286	Abk87286 Hepatitis

## ALIGNMENTS

RESULT 1  
AAD06793  
ID AAD06793 standard; DNA; 5882 BP.

XX AAD06793;

XX 06-AUG-2001 (first entry)

XX Plasmid pCMV-II-E2661-SAG.

XX Virus-like particle; immunogen; hepatitis B virus surface antigen; HBsAg;  
KW HBV; chimeric antigen; hepatitis C virus; HCV; virucide; vaccine;  
KW plasmid pCMV-II-E2661-SAG; cyclic; circular; ds.

XX Human cytomegalovirus.

XX Homo sapiens.

XX Hepatitis C virus.

XX Hepatitis B virus.

XX Bos sp.

XX Unidentified.

XX Chimeric.

Key	Location/Qualifiers
CDS	1992..3587
FT	/*tag= a
FT	/product= "HCV 561 E2 envelope glycoprotein-HBsAg S
FT	domain"
FT	1992..2900
FT	misc_feature
FT	/*tag= b









```

XX SQ Sequence 1914 BP; 349 A; 590 C; 535 G; 440 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,54e-110 Length: 1914
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 8 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x ACCT1824 (1-1914)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 604 GTGCTGCTGCTATTGGCGGCTC-----GAGCGGAAACCCACGTCACC 648

QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 649 GGGGGAAGTCCGCGCCACACTGTGCTCGGATTTGTTAGCCTCCTCGCACCAGGCGCAAG 708

QY 51 GluAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 709 CAGAACCTCAGCTGATCAACACACCGGAGTTGGCACCTCAATAGCAGGCCCTGAAC 768

QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
Db 769 TGAATGATAGCTCAACACCGGCTGTTGGCAGGCTTTTCTATCACCAAGTTCAAC 828

QY 91 SerSerGlyCysProGluAluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 829 TCTTCAGGCTGTCTGAGAGCTAGCCAGCTGCGACCCCTTACCGATTTTGACAGGCG 888

QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 889 TGGGCCCTATCAGTTATGCAACGGAAGCGGCCCGCCAGCAGGCCCTACTGCTGGCAC 948

QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 949 TACCCCAAAACCTTCGCGGTATTGTGCCCGCGAAGAGTGTGTGCTCGGTATTGTC 1008

QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1009 TTCACCTCCACGCCCCGTTGGTGGGAACGACGACAGGTGCGGCGCGCCACCTACGC 1068

QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyVal 190
Db 1069 TGGGTGAAATGATACGAGCGTCTTCTGCTTAAACAAATACAGCCCGCCGCTGGGCAAT 1128

QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1129 TGGTTCGGTTGTACCTGGATGAATCACTCACTGGATTACCAAGTGTGCGAGCGCCTCT 1188

QY 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 1189 TGTGTATCATCGAGGCGCGGCAACACACCTGCACTGCCCCACTGATTCCTTCGCGAAG 1248

QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 1249 CATCCGACGCCACATACTCTCGGTGGGCTCGGTCCTGGATCACACCGAGTGCCTG 1308

QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 1309 GTGACTACCGTATAGCTTTGGCATATCTTGTACCATCACTACATATATTAA 1368

QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaCysAsnTrpThrArg 290
Db 1369 ATCAGGATGTACGTGGAGGGTTCGAGCACAGGCTGGAAGCTGCTGCAACTGGAGCGG 1428

QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 1429 GCGCAAGTTGCGATCTGGAAGATAGGAGCAGGTCGAG 1467

```

```

RESULT 5
ADO00769
ID ADO00769 standard; DNA; 1914 BP.
XX AC ADO00769;
XX AC ADO00769;
XX DT 29-JUL-2004 (first entry)
XX XX HCV-1 E1/E2/p7 region nucleotide sequence SEQ ID NO:3.
XX DE HCV-1 E1/E2/p7 region nucleotide sequence SEQ ID NO:3.
XX KW fusion protein; hepatitis C virus; HCV; NS3; NS4; core polypeptide;
XX KW polypeptide; HCV-1; antiviral; vaccine; T cell activation;
XX KW E1/E2/p7 region; gene; ds; NS5.
XX OS Hepatitis C virus.
XX PH Key Location/Qualifiers
XX FT CDS 1..1914
XX FT /*tag= a
XX FT /product= "HCV-1 E1/E2/p7 region"
XX PN WO2004039950-A2.
XX XX 13-MAY-2004.
XX XX 24-OCT-2003; 2003WO-US033610.
XX XX 25-OCT-2002; 2002US-00281341.
XX PA (CHIR ) CHIRON CORP.
XX XX Houghton M, Coates S, Selby M, Paliard X;
XX WPI; 2004-376177/35.
XX DR P-PSDB; ADO00770.
XX PT New fusion protein comprises hepatitis C virus (HCV) polypeptides, which
XX PT consist of an NS3, an NS4, an NS4 and a core polypeptide of HCV, useful
XX PT for stimulating immune response, e.g. activating T cells of a vertebrate.
XX PS Disclosure; SEQ ID NO 3; 136pp; English.
XX CC The present invention describes a fusion protein comprising hepatitis C
XX CC virus (HCV) polypeptides, where the HCV polypeptides consist of an NS3,
XX CC an NS4, an NS5 and a core polypeptide of HCV, and the core polypeptide
XX CC consist of amino acids 1-121 of the HCV polypeptide, numbered relative to
XX CC the full-length HCV-1 polypeptide. Also described: (1) an isolated and
XX CC purified polynucleotide that encodes the fusion protein; (2) a
XX CC composition comprising the fusion protein or the isolated and purified
XX CC polynucleotide and a pharmaceutical excipient; and (3) activating T cells
XX CC of a vertebrate subject which recognise an epitope of an HCV polypeptide.
XX CC The fusion protein has antiviral activity, and can be used in vaccines.
XX CC The composition is useful for activating T cells of a vertebrate subject
XX CC which recognise an epitope of an of the NS3, NS4, NS5 and/or core
XX CC polypeptides. It is also useful for the manufacturing of a medicament for
XX CC activating T cells. The present sequence encodes an HCV-1 E1/E2/p7 region
XX CC amino acid sequence, which is used in the exemplification of the present
XX CC invention.
XX SQ Sequence 1914 BP; 349 A; 590 C; 535 G; 440 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,54e-110 Length: 1914
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 12 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x ADO00769 (1-1914)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30

```

```

Db      604 GTGCTGCTGCTATTTTCCGGCGTC-----GACGCGGAACCCACGTCAC 648
QY      31  GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
      |||
Db      649 GGGGGAAGTCCGGCCACACTGTGTCTGGATTGTTAGCTCTCGCACAGGCGCCAA 708
QY      51  GlnAsnValGlnLeuIleAsnThrAenGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
      |||
Db      709 CAGAACGTCCAGTGTATCAACACCAACGCGAGTTGCCACCTCAATAGACAGGCCCTGAAC 768
QY      71  CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
      |||
Db      769 TGCATGATAGCTCAACACCGGCTGTGTTGGCAGGGCTTTTCTATCACCACCAAGTTCAAC 828
QY      91  SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuLeuThrAspPheAspGlnGly 110
      |||
Db      829 TCTTCAGGCTGTCTCAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTTCACCAAGGC 888
QY      111 TrpGlyProIleSerTyrAlaAenGlySerGlyProAspGlnArgProTyrCysTrpHis 130
      |||
Db      889 TGGGGCCCTATCATGTTATGCCAACGGAAGCGGCCCGACGAGCGCCCTACTGCTGGCAC 948
QY      131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
      |||
Db      949 TACCCCCCAAACTTGGCGTATTGTGCCGCGAAGAGTGTGTGGTCCGGTATATTC 1008
QY      151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
      |||
Db      1009 TTCACCTCCAGGCCCTGTGTGGTGGAGACGACGACAGGTGGCGGCCGCCACCTACAGC 1068
QY      171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
      |||
Db      1069 TGGGGTGAAATATACGACGCTTCCTCGTCCTTAAACAATACACAGGCCACCGCTGGCAAT 1128
QY      191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
      |||
Db      1129 TGGTTCGGTGTACCTGGATGAACCTCACTGGATTCAACAAAGTGTGCGAGCGCTCCT 1188
QY      211 CysValIleGlyAlaGlyAlaAsnThrLeuHisCysProThrAspCysPheArgLys 230
      |||
Db      1189 TGTGTATCGGAGGGCGGCAACAACACCCCTGCACCTGCCACCTGATTTGCTTCGCAAG 1248
QY      231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
      |||
Db      1249 CATCCGGAGCCACATACATCTCGGTGGGCTCCGGTCCCTGGATCACACCGAGTGCCCTG 1308
QY      251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
      |||
Db      1309 GTCGACTACCCGTATAGGCTTTGGCATTTATCCTTGTACCATCAACTACACTATATTTAA 1368
QY      271 IleArgMetTyrValGlyValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
      |||
Db      1369 ATCAGGATGTACGTGGAGGGTCCGAGCAGGCTGGAAGCTGCCTGCAACTGGACGCG 1428
QY      291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
      |||
Db      1429 GGCGAACGTTGCCATCTGGAAGATAGGGACAGTCCGAG 1467

RESULT 6
AAT18246
ID      AAT18246 standard; DNA; 1989 BP.
AC
XX
AC      AAT18246;
XX
DT      27-AUG-2003 (revised)
DT      25-MAR-2003 (revised)
DT      04-JUN-1996 (first entry)
XX
XX      HCV1 E2/NS2 region DNA.
DE
XX      HCV1; E1 protein; E2 protein; truncation; vaccine; diagnosis; therapy;
KW      protein secretion; ds.
XX

```

```

OS      Hepatitis C virus.
XX
FH      Key      Location/Qualifiers
FT      sig_peptide 1..63
FT      /*tag= a
FT      mat_peptide 64..1986
FT      /*tag= b
XX
PN      WO9604301-A2.
XX
XX      15-FEB-1996.
XX
XX      26-JUL-1995; 95WO-US010035.
XX
XX      29-JUL-1994; 94US-00282959.
XX      25-JUL-1995; 95US-00506608.
XX
XX      (CHIR ) CHIRON CORP.
XX
XX      Selby M, Houghton M;
XX
XX      WPI; 1996-129331/13.
XX      P-P8DB; AAR92935.
XX
XX      Secretable hepatitis C virus E1 and E2 polypeptide(s) lacking all/part of
XX      the membrane spanning domain - useful in vaccines, and for diagnostic and
XX      therapeutic purposes, e.g. in assays for HCV.
XX
XX      Disclosure; Fig 4a-4c; 46pp; English.
XX
XX      The hepatitis c virus (HCV) E2/NS2 region (AAT18245) codes for the
XX      envelope E2+NS2 proteins (AAR92935). Elimination of sequences encoding
XX      the C-terminal membrane-spanning domain (approx. amino acids 337-361) of
XX      E2 provides a truncated protein that can be secreted when produced
XX      recombinantly in host cells. Truncated versions of the E1 protein (see
XX      AAR92934) are similarly produced. When co-expressed or combined after
XX      separate expression, the truncated E2 and E1 proteins form a complex. The
XX      truncated proteins and complex are useful in vaccines and for diagnostic
XX      and therapeutic apps. (Updated on 25-MAR-2003 to correct PR field.)
XX      (Updated on 27-AUG-2003 to correct OS field.)
XX
XX      Sequence 1989 BP; 373 A; 586 C; 568 G; 462 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1..61e-110      Length:      1989
Score:           1620.50      Matches:     283
Percent Similarity: 96.93%      Conservative: 1
Best Local Similarity: 96.59%      Mismatches:  4
Query Match:      92.81%      Indels:      5
DB:               2           Gaps:        1

US-10-715-665-7_COPY_1_303 (1-303) x AAT18246 (1-1989)

QY      11  ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
      |||
Db      31  GTGCTGCTGCTATTTTCCGGCGTC-----GACGCGGAACCCACGTCAC 75
      |||
QY      31  GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
      |||
Db      76  GGGGGAAGTCCGGCCACACTGTGTCTGGATTGTTAGCTCTCGCACAGGCGCCAA 135
      |||
QY      51  GlnAsnValGlnLeuIleAsnThrAenGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
      |||
Db      136 CAGAACGTCCAGTGTATCAACACCGGCTGTGCGAGGGCTTTTCTATCACCACCAAGTTCAAC 195
      |||
QY      71  CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
      |||
Db      196 TGCATGATAGCTCAACACCGGCTGTGCGAGGGCTTTTCTATCACCACCAAGTTCAAC 255
      |||
QY      91  SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuLeuThrAspPheAspGlnGly 110
      |||
Db      256 TCTTCAGGCTGTCTCAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTTCGACCAAGGC 315
      |||

```



QY 211 CysValIleGlyGlyAlaGlyValAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 DB 616 TGTGTCTATCGAGGGGGGGGCAACACACCTGCTGCACTGCCCACTGATTGCTTCGCAAG 675  
 QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProThrIleThrProArgCysLeu 250  
 DB 676 CATCGGACGCCACATACCTCTCGGTGCGGCTCCGCTCCGTCGATCACACCCAGGTGCGCTG 735  
 QY 251 ValAspTyrProTyrArgLeuThrHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
 DB 736 GTGACTACCCGTATAGGCTTTGGCATTATCTTGTACCATCACTACCATATTATAA 795  
 QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnThrThrArg 290  
 DB 796 ATCAGAGTGTAGTGGGAGGGTCCGAACACAGCTGGAAGCTGCTGCAACTGGAGCGG 855  
 QY 291 GlyIuArgCysAspLeuGluAspArgAspArgSerGlu 303  
 DB 856 GCGGAACGTTGCGATCTGGAAGATAGGAGACAGGTCCGAG 894  
 RESULT 8  
 ID AAQ05955 standard; cDNA; 8316 BP.  
 XX AC AAQ05955;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 23-JAN-1991 (first entry)  
 XX DE Hepatitis C virus open reading frame.  
 XX KW Hepatitis C virus; antiviral agent; ds.  
 XX OS Hepatitis C virus.  
 XX PN EP388232-A.  
 XX PD 19-SEP-1990.  
 XX PF 16-MAR-1990; 90EP-00302866.  
 XX PR 17-MAR-1989; 89US-00325338.  
 XX PR 20-APR-1989; 89US-00341334.  
 XX PR 18-MAY-1989; 89US-00355002.  
 XX PA (CHIR ) CHIRON CORP.  
 XX PI Houghton M, Choo QL, Kuo G;  
 XX WPI; 1990-284418/38.  
 XX DR P-PSDB; AAR08123.  
 XX PT Hepatitis C virus DNA - used for producing probes, polypeptide(s),  
 XX PT antibodies and anti-sense polynucleotide(s) for diagnosis and therapy.  
 XX PS Disclosure; Fig 16; 83pp; English.  
 XX CC HCV cDNA libraries were constructed using pooled serum from a chimpanzee  
 CC with chronic HCV infection. A lambda gt10 library was screened with  
 CC probes derived from previously isolated clones. The ORF is derived from  
 CC the overlapping clones p14a, CA167b, CA156e, CA84a, CA59a, K9-1, 12f,  
 CC 14i, 11b, 7f, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f,  
 CC 33g, 39c, 35f, 19g, 26g and 15e. Polypeptide encoded by this sequence can  
 CC be used to design probes for the detection of HCV nucleic acids, in  
 CC screening programmes for antiviral agents and in preparing blood free of  
 CC HCV. Antisense polynucleotides can be used to inhibit viral replication.  
 CC See also AAQ05956. (Updated on 25-MAR-2003 to correct PA field.)  
 XX SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;  
 Alignment Scores: 8.76e-110 Length: 8316  
 Pred. No.: 1620.50 Matches: 283  
 Score: XX

Percent Similarity: 96.93% Conservative: 1  
 Best Local Similarity: 96.59% Mismatches: 4  
 Query Match: 92.81% Indels: 5  
 DB: 2 Gaps: 1  
 US-10-715-665-7\_COPY\_1\_303 (1-303) x AAQ05955 (1-8316)  
 QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
 DB 778 GTGCTGCTGCTATTGTCGGGCTC-----GACGGGGAACCAACCGTCACC 822  
 QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
 DB 823 GGGGGAAGTCGGCCACACTGTGTGTGATTTGTTAGCTCCTCGCACAGGCCCAAG 882  
 QY 51 GluAsnValGluLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70  
 DB 883 CAGAACGTCAGCTGATCAACACCGCAGTGGCACCTCAATAGCACGGCCCTGAAAC 942  
 QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90  
 DB 943 TGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCAAGTTC AAC 1002  
 QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGly 110  
 DB 1003 TCTTCAGGCTGTCTCTGAGAGGCTAGCCAGCTGCCAGCCCTTACCGATTTTGACCGGCG 1062  
 QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130  
 DB 1063 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGCCAGCGGCCCTACTGTCTGGCAC 1122  
 QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
 DB 1123 TACCCCCCAAAACCTTGGGTATTGTGCCGGAAGAGTGTGTGTGTGTCGGTATATTGC 1182  
 QY 151 PheThrProSerProValValGlyThrThrAspArgSerGlyValAlaProThrTyrSer 170  
 DB 1183 TTCTCTCCAGCCCGCTGTGTGTGGAAACGACCGCAGGTGCGGGCGGCCACCTACAGC 1242  
 QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
 DB 1243 TGGGGTGAAATGATACGACGCTCTTCGTTTAAACATACACAGGCCACCGCTGGCAAT 1302  
 QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
 DB 1303 TGGTTTGGTTGTACTGCGATGAACCTCAACTGGATTTCACAAAGTGTGCGAGCGCTCCT 1362  
 QY 211 CysValIleGlyGlyAlaGlyAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 DB 1363 TGTGTCTATCGGAGGGGGGGCAACACACCTGTCACCTGCCCCCAGTATTGCTTCGCAAG 1422  
 QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
 DB 1423 CATCGGACGCCACATACCTCTCGGTGCGGCTCCGTCCTGGATCACACCCAGGTGCGCTG 1482  
 QY 251 ValAspTyrProTyrArgLeuThrHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
 DB 1483 GTGACTACCCGTATAGGCTTTGGCATTATCTTGTACCATCAACTACACCATATTATAA 1542  
 QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnThrThrArg 290  
 DB 1543 ATCAGAGTGTAGTGGGAGGGTCCGAACACAGGCTGGAAGCTGCTGCAACTGGAGCGCG 1602  
 QY 291 GlyIuArgCysAspLeuGluAspArgAspArgSerGlu 303  
 DB 1603 GCGGAACGTTGCGATCTGGAAGATAGGAGACAGGTCCGAG 1641  
 RESULT 9  
 ID AAA75296 standard; cDNA; 8316 BP.  
 XX AC AAA75296;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 23-JAN-1991 (first entry)  
 XX DE Hepatitis C virus open reading frame.  
 XX KW Hepatitis C virus; antiviral agent; ds.  
 XX OS Hepatitis C virus.  
 XX PN EP388232-A.  
 XX PD 19-SEP-1990.  
 XX PF 16-MAR-1990; 90EP-00302866.  
 XX PR 17-MAR-1989; 89US-00325338.  
 XX PR 20-APR-1989; 89US-00341334.  
 XX PR 18-MAY-1989; 89US-00355002.  
 XX PA (CHIR ) CHIRON CORP.  
 XX PI Houghton M, Choo QL, Kuo G;  
 XX WPI; 1990-284418/38.  
 XX DR P-PSDB; AAR08123.  
 XX PT Hepatitis C virus DNA - used for producing probes, polypeptide(s),  
 XX PT antibodies and anti-sense polynucleotide(s) for diagnosis and therapy.  
 XX PS Disclosure; Fig 16; 83pp; English.  
 XX CC HCV cDNA libraries were constructed using pooled serum from a chimpanzee  
 CC with chronic HCV infection. A lambda gt10 library was screened with  
 CC probes derived from previously isolated clones. The ORF is derived from  
 CC the overlapping clones p14a, CA167b, CA156e, CA84a, CA59a, K9-1, 12f,  
 CC 14i, 11b, 7f, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f,  
 CC 33g, 39c, 35f, 19g, 26g and 15e. Polypeptide encoded by this sequence can  
 CC be used to design probes for the detection of HCV nucleic acids, in  
 CC screening programmes for antiviral agents and in preparing blood free of  
 CC HCV. Antisense polynucleotides can be used to inhibit viral replication.  
 CC See also AAQ05956. (Updated on 25-MAR-2003 to correct PA field.)  
 XX SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;  
 Alignment Scores: 8.76e-110 Length: 8316  
 Pred. No.: 1620.50 Matches: 283  
 Score: XX



PR 16-MAR-1990; 90EP-00302866.  
 XX (CHIR ) CHIRON CORP.  
 PA Houghton M, Choo Q, Kuo G;  
 XX WPI; 2004-193149/19.  
 DR P-PSDB; ADN35976.  
 XX Novel purified hepatitis C virus polypeptide comprising epitope encoded  
 PT by hepatitis C virus cDNA, useful as vaccine for treating hepatitis C  
 PT virus.  
 XX Example 1; Fig 16; 79pp; English.  
 XX The present invention relates to hepatitis C virus (HCV) proteins and  
 CC cDNA sequences. The sequences are useful in immunoassays for detecting  
 CC antibodies directed against HCV antigen; preparing host cells transformed  
 CC with a recombinant polynucleotide; screening antiviral agents and  
 CC determining the effect of antiviral agent in inhibiting viral replication  
 CC in cell culture system; and developing vaccine for treating HCV  
 CC infection.  
 XX Sequence 8316 BP; 1671 A; 2537 C; 2338 G; 1770 T; 0 U; 0 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 8,76e-110 Length: 8316  
 Score: 1620.50 Matches: 283  
 Percent Similarity: 96.93% Conservative: 1  
 Best Local Similarity: 96.59% Mismatches: 4  
 Query Match: 92.81% Indels: 5  
 DB: 12 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x ADN35977 (1-8316)

QY 11 ValLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
 DB 778 GTGCTGCTGCTATTTCGCCGGCGTC-----GACGCGGAACCCACGTCACC 822

QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
 DB 823 GGGGGAAGTCCGGCCACACTGTGCTGATTGTTAGCTCTCTCGCACCGGCCAAG 882

QY 51 GlnAsnValGlnLeuLeuAlaThrAenGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70  
 DB 883 CAGAACGTCAGCTGATCAACACCAACGCGCAGTGGCAGCTCAATAGCACGCGCCCTGAAC 942

QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrrHisHisPheAsn 90  
 DB 943 TGCATATGATAGCTCAACACCGGCTGTTGGCAGGCGCTTTCTATCACCACCAAGTTCAAC 1002

QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAepPheAepGlnGly 110  
 DB 1003 TCTTCAGGCTGTCTCAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTTGACACAGGCG 1062

QY 111 TrpGlyProLeuSerTrpAlaAlaGlySerGlyProAspGlnArgProTyrrCysTrpHis 130  
 DB 1063 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACGAGCGGCCCTACTGCTGGCAC 1122

QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrrCys 150  
 DB 1123 TACCCCCCAAACTTGGCGGTATTGTCCCGCGGAAGAGTGTGTGGTTCGGTATATTGC 1182

QY 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrrSer 170  
 DB 1183 TTCACCTCCAGCCCGTGGTGGGGAACGACGACAGGTTCGGCGCGGCCACCTACAGC 1242

QY 171 TrpGlyGlnAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
 DB 1243 TGGGGTGAATAATGATACGGACGCTTTCGTCCTTAAACAATACCAAGGCCACGCTGGGCAAT 1302

QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210

Db 1303 TGGTTCGGTTGTACTGATGAACCTCAACTGATTCACAAAGTGTGCGGAGCGCTCT 1362  
 QY 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 Db 1363 TGTGTCACTCGGAGGCGGCAACCAACCCCTGCACTGCCCACTGATTGCTTCGCAAG 1422  
 QY 231 HisProAspAlaThrTyrrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
 Db 1423 CATCCGACGCGCACATCTCTCGGTGCGGCTCCCGTCCCTGGATCACACCCAGGTGCTG 1482  
 QY 251 ValAspTyrrProTyrrArgLeuTrpHisTyrrProCysThrIleAsnTyrrThrIlePheLys 270  
 Db 1483 GTCGACTACCCGTATAGGCTTTGGCATTTATCTCTGTACCATCACTACACATATTTAA 1542  
 QY 271 IleArgMetTyrrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
 Db 1543 ATCAGATGTAGTGGGAGGGGTGCAACACAGGCTGGAAGCTGCCTGCAACTGGACGCGG 1602  
 QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
 Db 1603 GCGCAACGTTGCGATCTGGAACAGACAGGACAGGTCCGAG 1641

RESULT 11  
 AAQ05956  
 ID AAQ05956 standard; DNA; 9185 BP.  
 XX  
 AC AAQ05956;  
 XX 25-MAR-2003 (revised)  
 DT 23-JAN-1991 (first entry)  
 XX  
 DE Sense strand of the compiled Hepatitis C virus cDNA sequence.  
 XX  
 KW Hepatitis C virus (HCV); antiviral agent; ss.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_RNA 1..1667  
 FT /tag= b  
 FT /note= "epitope within this region is claimed"  
 FT CDS 320..9185  
 FT /tag= a  
 FT misc\_RNA 8978..9185  
 FT /tag= c  
 FT /note= "encodes an epitope that is claimed"  
 XX  
 PN EP388232-A.  
 XX  
 PD 19-SEP-1990.  
 XX  
 PF 16-MAR-1990; 90EP-00302866.  
 XX  
 PR 17-MAR-1989; 89US-00325338.  
 PR 20-APR-1989; 89US-00341334.  
 PR 18-MAY-1989; 89US-00355002.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Houghton M, Choo QL, Kuo G;  
 XX  
 DR WPI: 1990-284418/38.  
 DR P-PSDB; AAR08124.  
 XX  
 PT Hepatitis C virus DNA - used for producing probes, polypeptide(s),  
 PT antibodies and anti-sense polynucleotide(s) for diagnosis and therapy.  
 XX  
 PS Disclosure; Fig 17; 83pp; English.  
 XX  
 CC HCV cDNA libraries were constructed using pooled serum from a chimpanzee  
 CC with chronic HCV infection. A lambda gtl1 library was screened with  
 CC probes derived from previously isolated clones. The ORF is derived from  
 CC the overlapping clones b114a, ag30a, CA205a, CA216a, p14a,



CC CA167b, CA156e, CA84a, CA59a, K9-1, 26j, 13i, 12f, 13i, 12f, 14i, 11b, 7f, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g, 39c, 35f, 19g, 26g, 15e, b5a and 16jh. These clones extend the sequence of the HCV genome reported in EP-318216. The upstream region from nucleotides -319 to +1348 (-1-1667 in this file) is covered by clones b114a, 18g, ag30a, CA205a, CA290a, CA216a, p14a, CA167b, CA156e, CA84a and CA59a; nucleotides 8659-8866 (-8978-9185 in this file) are covered by clones b5a and 16jh. See also AAQ05955. (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9.86e-110 Length: 9185  
 Score: 1620.50 Matches: 283  
 Percent Similarity: 96.93% Conservativeness: 1  
 Best Local Similarity: 96.59% Mismatches: 4  
 Query Match: 92.81% Indels: 5  
 DB: 2 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x AAQ05956 (1-9185)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
 DB 1439 GTGCTGCTGCTATTGGCGGCGTC-----GACGCGGAACCCACGTCACC 1483  
 QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
 DB 1484 GGGGGAAGTCCGCGCCACACTGTGCTGTGATTGTTAGCTCTCTCGCAGGCGCCAG 1543  
 QY 51 GlnAsnValGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 70  
 DB 1544 CAGAACGTCAGCTGATCAACACACGCGGAGTTGGCACTCAATAGCAGGCGCTGAAC 1603  
 QY 71 CysAsnAspSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 90  
 DB 1604 TGCATGATAGCTTCAACACCGGCTGGTGGAGGGCTTTTCTATACCAAGTTCAAC 1663  
 QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
 DB 1664 TCTTCAGCTGTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTTTGACGAGGC 1723  
 QY 111 TrpGlyProLeuSerTyAlaAsnGlySerGlyProAspGlnArgProTyTrpHis 130  
 DB 1724 TGGGGCCCTATCATGTTATGCAACGGAAGCGGCGCCGACGAGCGGCGCTACTGCTGCAC 1783  
 QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyCys 150  
 DB 1784 TACCCCAAAACCTTTCGGTATTGTGCGCGGGAAGAGTGTGTGTCGGTATTATTC 1843  
 QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTySer 170  
 DB 1844 TTCACCTCCAGCCCGTGGTGGTGGNACCGACGAGTTCGGGCGGCCACCTACAGC 1903  
 QY 171 TrpGlyLeuAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
 DB 1904 TGGGGTGAAATGATACGAGCTCTTCGCTCTTAAACAATACGAGGCGCCGCTGGCAAT 1963  
 QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
 DB 1964 TGGTTCGGTGTACCTGGATGAACACTCACTGGGATTCACCAAGTGTGCGGAGCGCTCCT 2023  
 QY 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 DB 2024 TGTGTCATCGAGGGGGGCGACACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2083  
 QY 231 HisProAspAlaThrTySerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
 DB 2084 CATCCGAGCCACATACTCTCGGTGCGGCTCGGTCTCGGTCTCGGTCTCGGTCTCGGTCT 2143  
 QY 251 ValAspTyProTyArgLeuTrpHisTyProCysThrIleAsnTyThrIlePheLys 270  
 DB 2144 GTGACTTACCCGTATAGGCTTTGGCATTTATCTTTGATCCATCACTACCATATTTAA 2203

QY 271 IleArgMetTyValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
 DB 2204 ATCAGGATGTACGTGGAGGGGTGGAACACAGAGCTGGAAGCTGCTGCACTGGACCGG 2263  
 QY 291 GlyGluArgCysAspLeuGluAlaAspArgAspArgSerGlu 303  
 DB 2264 GGGCAACGTTGCGATCTGGAAGACAGGACAGGAGTCCGAG 2302

# RESULT 12

AAQ10566  
 ID AAQ10566 standard; DNA; 9185 BP.

XX AC AAQ10566;

XX XX 25-MAR-2003 (revised)

DT 29-APR-1991 (first entry)

XX XX Hepatitis C virus strain 1 DNA.

XX KW Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;

XX KW viral infections; ss.

OS Hepatitis C virus.

XX PN EP414475-A.

XX PD 27-FEB-1991.

XX PF 21-AUG-1990; 90BP-00309120.

XX PR 25-AUG-1989; 89US-00398667.

XX PA (CHIR ) CHIRON CORP.

XX PI Weiner AJ, Steimer KS;

XX DR WPI; 1991-059670/09.

XX PT Cell lines infected with hepatitis C virus - are used as source of  
 PT antigens for detection of HCV antibodies, for vaccines, and for screening  
 PT anti-viral agents.

XX PS Disclosure; Fig 1; 24pp; English.

XX CC This is a hepatitis C virus (HCV) composite cDNA sequence, deduced using  
 CC overlapping clones. a compsn. contg. the antigenic protein encoded by  
 CC this sequence is useful for detecting anti-HCV anti- bodies (Abs) and for  
 CC screening an agent which inhibits HCV replic- ation. A cell line infected  
 CC with this virus can be used as a source of antigens. The antigen is  
 CC useful for preparing vaccines for treating viral infections. See also  
 CC AAQ10567. (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 9.86e-110 Length: 9185  
 Score: 1620.50 Matches: 283  
 Percent Similarity: 96.93% Conservativeness: 1  
 Best Local Similarity: 96.59% Mismatches: 4  
 Query Match: 92.81% Indels: 5  
 DB: 2 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x AAQ10566 (1-9185)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
 DB 1439 GTGCTGCTGCTATTGGCGGCGTC-----GACGCGGAACCCACGTCACC 1483  
 QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
 DB 1484 GGGGGAAGTCCGCGCCACACTGTGCTGTGATTGTTAGCTCTCTCGCAGGCGCAAG 1543  
 QY 51 GlnAsnValGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 70

```
Db      1544 CAGACGTCAGCTGATCAACACCAACGCGAGTTGGCACTCAATAGACACGCGCTTGAC 1603
Qy      71  CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyRHisHisLysPheAsn 90
Db      1604 TGCATATGATAGCTCAACACCGCTGGTTGGCAGGCGCTTTCTATCACCACCAAGTTCAAC 1663
Qy      91  SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db      1664 TCTTCAGGCTGTCTCAGAGGCTAGCCAGCTGCCAGCCCTTACCAGATTTTACCAGGCG 1723
Qy      111 TrpGlyProIleSerTyRAlaAsnGlySerGlyProAspGlnArgProTyRHis 130
Db      1724 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACAGCGCCCTTACTGTGGCAC 1783
Qy      131 TyRProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyRHis 150
Db      1784 TACCCCCCAAAACCTTGGCGTATTTGTCGCCGGAAGAGTGTGTGGTCCGGTATATTGC 1843
Qy      151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyRHis 170
Db      1844 TTCACTCCAGCGCCGTGTGGTGGGAACGACGACAGTGTGGCGCGCCCACTACAGC 1903
Qy      171 TrpGlyGluAsnAspThrAspValPheValLeuAsnThrArgProProLeuGlyAsn 190
Db      1904 TGGGGTGAATATGATACGAGCTTTCGTCTTAACATACAGGCCACCGTGGCAAT 1963
Qy      191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db      1964 TGTTCCGTTGTACCTGGATGAATCAACTGGATTCAACAAAGTGTGGCGAGCGCTCCT 2023
Qy      211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db      2024 TGTGTATCGAGGGCGGGCAACAACACCTCGACTGCCCTCACTGATTTGCTTCGCAAG 2083
Qy      231 HisProAspAlaThrTyRSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db      2084 CATCCGAGCCACATACTCTCGTGGGCTCGGTCTCGGTCTCGATCAACCCAGGTGCTG 2143
Qy      251 ValAspTyRProTyRArgLeuTrpHisTyRProCysThrIleAsnTyRThrlePheLys 270
Db      2144 GTCGACTACCGGTATAGGCTTTGGCATTTATCTGTGTACCATCAACTACACCATATTTAA 2203
Qy      271 IleArgMetTyRValGlyValGluHisArgLeuGluAlaAlaCysAsnThrTrpArg 290
Db      2204 ATCAGGATGTACGTGGAGGGGTGGAACACAGCTGGAAGCTGCACTGCACTGGACGG 2263
Qy      291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db      2264 GCGAAGCTTGGCATCTGGAAGACAGGACAGGTCCGAG 2302

RESULT 13
ID AAX00459
XX AAX00459 standard; DNA; 9185 BP.
XX AAX00459;
XX AAX00459;
DT 10-MAY-1999 (first entry)
XX
DE Hepatitis C virus 1 ORF1.
XX
KW HCV; non-A, non-B hepatitis; NANBH; diagnosis; vaccine; antigen;
KW antibody; immunoassay; assay; SB.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 320..9185
FT /tag= a
XX
XX US5856437-A.
XX
PD 05-JAN-1999.
```

```
XX      03-NOV-1994; 94US-00334255.
PF
XX      15-SEP-1989; 89US-00408045.
PR      21-DEC-1989; 89US-00456142.
PR      04-JAN-1991; 91US-00637380.
PR      02-AUG-1993; 93US-00101280.
PR      24-FEB-1994; 94US-00201066.
XX
XX (NAHE-) NAT INST OF HEALTH JAPAN.
PA (CHIR) CHIRON CORP.
XX
XX Weiner AJ, Miyamura T, Irvine BD, Kolberg JA, Han J, Saito I;
PI Cha T, Houghton M;
XX
XX WPI; 1999-105191/09.
XX
XX Antigenic polypeptides from J1 and J7 hepatitis C virus isolates - useful
PT as immunoassay reagents, for raising antibodies and as vaccine
PT components.
XX
XX Disclosure; Fig 12A-C; 44pp; English.
XX
XX This is the nucleotide sequence of ORF1 of hepatitis C virus prototype
CC HCV-1. The invention relates to 2 new isolates, J1 and J7, of HCV. These
CC new isolates comprise nucleotide (see AAX00400-04) and amino acid (see
CC AAX0583-87) sequences which are distinct from HCV1. Thus J1 and J7
CC provide new polynucleotides and polypeptides useful in diagnostics,
CC recombinant protein production and vaccine development. Claimed
CC polypeptides can be used: (i) as immunoassay reagents, or standards, to
CC detect HCV antibodies, e.g. for diagnosing infection or screening donated
CC blood; (ii) to generate specific antibodies (used for detecting the
CC corresponding polypeptide, to screen for antiviral agents, for virus
CC isolation and for passive immunisation); (iii) in protective or
CC therapeutic vaccines, and (iv) for isolation of non-A, non-B viruses
XX
XX Sequence 9185 BP; 1849 A; 2793 C; 2605 G; 1938 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 9,86e-110 Length: 9185
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 2 Gaps: 1
US-10-715-665-7_COPY_1_303 (1-303) x AAX00459 (1-9185)
Qy      11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db      1439 GTGCTGTCTATTTCGCGGCTC-----GACGCGGAACCCACGTCACC 1483
Qy      31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db      1484 GGGGGAAGTGGCGGCACACTGTGTCTGGATTGTGTAGCTCTCGCACAGCGCGCAAG 1543
Qy      51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db      1544 CAGAACGTCAGCTATCAACACCGCTGGTTGGCAGGCGCTTTCTATCACCACCAAGTTCAAC 1603
Qy      71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyRHisHisLysPheAsn 90
Db      1604 TGCATATGATAGCTCAACACCGCTGGTTGGCAGGCGCTTTCTATCACCACCAAGTTCAAC 1663
Qy      91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db      1664 TCTTCAGGCTGTCTCAGAGGCTAGCCAGCTGCCAGCCCTTACCAGATTTTACCAGGCG 1723
Qy      111 TrpGlyProIleSerTyRAlaAsnGlySerGlyProAspGlnArgProTyRHis 130
Db      1724 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACAGCGCCCTTACTGTGGCAC 1783
Qy      131 TyRProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyRHis 150
```

```
Db 1784 TACCCCAACACCTTCGGTATTTGTCGCCGAGAGTGTGTGTGGTTCGGTATATTC 1843
Qy 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1844 TTCACCTCCAGCCCCGTTGGTGGGAAACGACACAGGTGGGGCGGCCACCTACAGC 1903
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1904 TGGGGTGAAATGATACGGAGCTCTTCGTCTTAACAATACACAGCCACCGCTGGGCAAT 1963
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1964 TGGTTCGGTGTACCTCGATGAACCTCAACTGGATTACCAAGTGTGGGAGCGCTTCCT 2023
Qy 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 2024 TGTGTCACTCGAGGGGGGGGCAACACCTTCGACCTGCCCACTGATTCCTCCGAG 2083
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 2084 CATCCGACGCCACATACCTCGGTGGGCTCCGTCCTGGATCACCACCGAGTGCCTG 2143
Qy 251 ValAspTyrProTyrArgLeuThrHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 2144 GTCGACTACCGGTATAGGCTTTGGCATTTATCTTGTACCATCAACTACACCATATTTAA 2203
Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
Db 2204 ATCAGGATGTACGTGGAGGGGTGCAACACAGGCTGGAGCTGCTGCAACTGACGCG 2263
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 2264 GCGCAACGTTGCGATCTGGAAGACAGGACAGGTCGAG 2302

RESULT 14
AAZ26737
ID AAZ26737 standard; DNA; 9185 BP.
XX
AC AAZ26737;
XX
DT 21-JUN-1999 (first entry)
XX
DE Nucleotide sequence of the ORF of HCV1.
XX
KW HCV; J7 isolate; J1 isolate; HCV1; immunoassay; asiatic strain;
KW diagnosis; HCV infection; blood screening; immunisation; antiviral; ss.
XX
OS Hepatitis C virus.
XX
PN US5871903-A.
XX
PD 16-FEB-1999.
XX
PF 08-MAY-1995; 95US-00436965.
XX
PR 15-SEP-1989; 89US-00408045.
PR 21-DEC-1989; 89US-00456142.
PR 04-JAN-1991; 91US-00637380.
PR 02-AUG-1993; 93US-00101280.
PR 24-FEB-1994; 94US-00201066.
PR 03-NOV-1994; 94US-00334255.
XX
(NAHE-) NAT INST OF HEALTH JAPAN.
PA (CHIR ) CHIRON CORP.
XX
PI Saito I, Miyamura T;
XX
WIPI; 1999-166619/14.
XX
DR Immunoassays for Asiatic strains of hepatitis C virus - for diagnosis of
PT infection and screening blood supplies.
XX
```

PS Disclosure; Fig 12A-C; 43pp; English.

XX The present sequence represents the sequence of ORF1 of hepatitis C virus (HCV) isolate HCV1. The specification describes new isolates of HCV, J1 and J7. J1 and J7 isolates comprise sequences which are distinct from the prototype HCV isolates, HCV1. The specification describes immunoassays for HCV based on antigens from Asiatic strains not cross-reactive with HCV-1. The assays are used for diagnosis of HCV infection and to screen donated blood. The anti-HCV antibodies are also useful therapeutically and prophylactically (passive immunisation); in screening for antiviral agents; for isolation, purification and identification of non-A, non-B hepatitis virus (e.g. by affinity chromatography) and to raise anti-idiotypic antibodies (useful for treatment or diagnosis and to determine immunogenic regions of the HCV antigens)

XX SQ Sequence 9185 BP; 1849 A; 2792 C; 2606 G; 1938 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 9,86e-110 Length: 9185  
Score: 1620.50 Matches: 283  
Percent Similarity: 96.93% Conservative: 1  
Best Local Similarity: 96.59% Mismatches: 4  
Query Match: 92.81% Indels: 5  
DB: 2 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x AAZ26737 (1-9185)

```
Qy 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1439 GTGCTGCTGCTATTTCGCCGGCGTC-----GACGGGAAACCCACGTCAAC 1483
Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1484 GGGGGAAGTCCCGCCACACTGTGTCTGGATTGTTAGCTCTTCGACACAGCGCCAG 1543
Qy 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 1544 CAGAACGTCCAGCTGATCAACACACCGGCGAGTTGGCACCCTCAATAGCAGCGCCTCAAC 1603
Qy 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
Db 1604 TGCAATGATAGCTCAACACCGGCTGGTGGGAGGGCTTTTCTATCACCACAAAGTTCAAC 1663
Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1664 TCTTCAGGCTGCTCCTCAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTTTGACGAGGC 1723
Qy 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 1724 TGGGGCCCTATCATGTTATGCCAACGGAAGCGGCGCCGACCCAGCGGCCCTACTGCTGGCAC 1783
Qy 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 1784 TACCCCCCAAAACCTTGGCGTATTGTGCCCGGGAAGAGTGTGTGTCCGGTATATATTC 1843
Qy 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1844 TTCACCTCCAGCCCCGTTGGTGGGAAACGACACAGGTCGGGGCGGCCACCTACAGC 1903
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1904 TGGGGTGAAATGATACGGAGCTCTTCGTCTTAACAATACACAGCGCCACCGCTGGGCAAT 1963
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1964 TGGTTCGGTGTACCTCGATGAACCTCAACTGGATTACCAAGTGTGGGAGCGCTTCCT 2023
Qy 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 2024 TGTGTCACTCGAGGGGGGGGCAACACCTTCGACCTGCCCACTGATTCCTCCGAG 2083
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
```

Db 2084 CATCGGAGCCACATACCTCTCGGTCCGGCTCGGTCCCTGGATCACACCCAGTGCCTG 2143  
 QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
 Db 2144 GTCGACTACCCGTATAGGCTTTGGCATTATCTCTGTACCATCAACTACACCATATTTAAA 2203  
 QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
 Db 2204 ATCAGGATGACGTGGGGGGGTGGAACACAGGCTGGAAAGCTGCCCTGCAACTGGAGCGG 2263  
 QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
 Db 2264 GCGGAACGTTGGATCTGGAACACAGGACAGGTCGGAG 2302

RESULT 15

ADF66068

ID ADF66068 standard; DNA; 9185 BP.

XX AC ADF66068;

XX 12-FEB-2004 (first entry)

XX Hepatitis C virus 1 ORF nucleotide sequence.

XX Hepatitis C virus; HCV; screening; identification; non A non B hepatitis;

XX NANBH; gene; ds.

XX Hepatitis C virus.

XX US959092-A.

XX 28-SEP-1999.

XX 08-MAY-1995; 95US-00436966.

XX 15-SEP-1989; 89US-00408045.

XX 21-DEC-1989; 89US-00456142.

XX 04-JAN-1991; 91US-00637380.

XX 02-AUG-1993; 93US-00101280.

XX 24-FEB-1994; 94US-00201066.

XX 03-NOV-1994; 94US-00334255.

XX (CHIR ) CHIRON CORP.

XX (NAHE-) NAT INST OF HEALTH JAPAN.

XX Saito I, Miyamura T;

XX WPI; 1999-570774/48.

XX Polynucleotide probes useful for screening and identifying carriers of

XX non A non B hepatitis.

XX Disclosure; Fig 12; 44pp; English.

XX The present invention describes polynucleotides sequences from Hepatitis

XX C virus (HCV) isolates J7 C/E, J1 E, J1 E/NS1, J1 NS3, and J1 NS5 or from

XX deposits BP-2593, BP-2594, BP-2637 and BP-2638. The

XX polynucleotide sequences are useful as probes for screening and

XX identifying carriers of non A non B hepatitis (NANBH). The present

XX sequence is used in the exemplification of the present invention.

XX Sequence 9185 BP; 1849 A; 2792 C; 2606 G; 1938 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9,86e-110	Length:	9185
Score:	1620.50	Matches:	283
Percent Similarity:	96.93%	Conservative:	1
Best Local Similarity:	96.59%	Mismatches:	5
Query Match:	92.81%	Indels:	5
DB:	2	Gaps:	1

US-10-715-665-7\_COPY\_1\_303 (1-303) x ADF66068 (1-9185)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
 Db 1439 GTGCTGCTGCTATTTTGGCCGGGTC-----GAGCGGAAACCCACGTCAC 1483  
 QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
 Db 1484 GGGGGAAGTGGCGGCACACCTGTGCTCGATTGTTAGCCTCTCGCACCAGCGCCAAG 1543  
 QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70  
 Db 1544 CAGAACGTCACGCTCATCAACACCAACGCGAGTTGGCACCTCAATAGCACGCGCCCTGAAC 1603  
 QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisIleLysPheAsn 90  
 Db 1604 TGCAATGATAGCCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAAGTTCAAC 1663  
 QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
 Db 1664 TCTTCAGGCTGCTCTGAGAGGCTAGCCAGCTGCCAGCCCTTACCGATTTTGACCCAGGCG 1723  
 QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130  
 Db 1724 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACGACGCGCCCTTACTGTGGCAC 1783  
 QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
 Db 1784 TACCCCCCAAAACCTTGGCGGTATTGTCCCGCGAAGAGTGTGTGGTCCGGTATATTGC 1843  
 QY 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170  
 Db 1844 TTCACCTCCAGCCCGCTGGTGGGAACGACCGACAGTGGCGCGCCACCTACACG 1903  
 QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
 Db 1904 TGGGGTGAATAATGATACGACGCTCTTCGTCTTAAACATACACGCGCCACCGTGGGCAAT 1963  
 QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
 Db 1964 TGGTTCCGGTTGTACCTGGATGAACCTCAACTGGATTCCAAAGTGTGGGAGCGCTCCT 2023  
 QY 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 Db 2024 TGTGTCAATCGAGGGCGGGCAACACACCTCGCTGCCTGCCCTGCTGCTTCCGCAAG 2083  
 QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
 Db 2084 CATCGGAGCGCACATACCTCTCGGTGGGCTCCGGTCCCTGGATCACACCCAGGTGCTG 2143  
 QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
 Db 2144 GTCGACTACCCGTATAGGCTTTGGCATTATTCCTTGTACCATCAACTACACCATATTTAAA 2203  
 QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
 Db 2204 ATCAGGATGATCGTGGAGGGGTGCAACACAGGCTGGAAAGCTGCCCTGCAACTGGAGCGG 2263  
 QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
 Db 2264 GCGGAACGTTGGGATCTGGGAACAGGAGGACAGGTCGGAG 2302

Search completed: April 15, 2005, 06:42:44

Job time : 633 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 15, 2005, 04:52:33 ; Search time 202 Seconds  
(without alignments)

2454.415 Million cell updates/sec

Title: US-10-715-665-7\_COPY\_1\_303

Perfect score: 1746

Sequence: 1 MDAMKRLGCVLLCGAVFV.....AACNWTGRCDLDRDRSE 303

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10715665/runat.13042005.165824.28194/app.query.fasta.1.455  
-DB=Issued PatentsNA -OFMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LIST=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-MODE=LOCAL -OUTFMT=ptc -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15  
-USER=US10715665 @CGN 1.1 69 @runat.13042005.165824.28194 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1746	100.0	5882	4	US-09-721-480-6
2	1620.5	92.8	1207	1	US-08-460-806-12
3	1620.5	92.8	1207	1	US-08-325-630-12
4	1620.5	92.8	1210	2	US-08-483-695-35
5	1620.5	92.8	1210	2	US-07-965-285-35
6	1620.5	92.8	1210	2	US-08-487-231-35
7	1620.5	92.8	1210	3	US-09-201-912-35
8	1620.5	92.8	1989	3	US-08-824-057-4
9	1620.5	92.8	1989	3	US-09-415-582-4
10	1620.5	92.8	1989	4	US-09-693-596-3
11	1620.5	92.8	8316	3	US-08-444-818-88
12	1620.5	92.8	8987	3	US-08-444-818-137

13	1620.5	92.8	9185	3	US-08-444-818-122	Sequence 122, App
14	1620.5	92.8	9185	3	US-08-444-818-123	Sequence 123, App
15	1620.5	92.8	9379	3	US-09-388-874-1	Sequence 1, Appli
16	1620.5	92.8	9379	4	US-09-916-359-1	Sequence 1, Appli
17	1620.5	92.8	9401	1	US-07-910-760-9	Sequence 9, Appli
18	1620.5	92.8	9401	1	US-08-440-519-9	Sequence 9, Appli
19	1620.5	92.8	9401	3	US-08-440-549-9	Sequence 9, Appli
20	1620.5	92.8	9401	3	US-08-823-895A-25	Sequence 25, Appli
21	1615.5	92.5	9379	3	US-08-444-818-176	Sequence 176, App
22	1615.5	92.5	9401	5	PCT-US91-02225-9	Sequence 9, Appli
23	1526	87.4	5125	1	US-08-453-552-11	Sequence 11, Appli
24	1526	87.4	5125	2	US-08-710-637-11	Sequence 11, Appli
25	1526	87.4	5125	5	PCT-US93-00907-11	Sequence 11, Appli
26	1526	87.4	5323	1	US-08-453-552-9	Sequence 9, Appli
27	1526	87.4	5323	2	US-08-710-637-9	Sequence 9, Appli
28	1526	87.4	5323	5	PCT-US93-00907-9	Sequence 9, Appli
29	1521.5	87.1	9401	2	US-08-432-693-1	Sequence 1, Appli
30	1521.5	87.1	9416	3	US-08-811-566-19	Sequence 19, Appli
31	1521.5	87.1	9416	3	US-09-034-756-19	Sequence 19, Appli
32	1518.5	87.0	9365	4	US-09-827-688-7	Sequence 7, Appli
33	1515.5	86.8	9646	3	US-08-811-566-1	Sequence 1, Appli
34	1515.5	86.8	9646	3	US-09-034-756-1	Sequence 1, Appli
35	1515.5	86.8	12980	3	US-08-811-566-5	Sequence 5, Appli
36	1515.5	86.8	12980	3	US-09-034-756-5	Sequence 5, Appli
37	1514	86.7	7106	1	US-08-188-281B-8	Sequence 8, Appli
38	1514	86.7	7106	1	US-08-453-552-5	Sequence 5, Appli
39	1514	86.7	7106	2	US-08-710-637-5	Sequence 5, Appli
40	1514	86.7	7106	5	PCT-US93-00907-5	Sequence 5, Appli
41	1514	86.7	7106	5	PCT-US94-07280-8	Sequence 8, Appli
42	1514	86.7	7106	5	PCT-US95-01087-8	Sequence 8, Appli
43	1514	86.7	7298	1	US-08-453-552-3	Sequence 3, Appli
44	1514	86.7	7298	2	US-08-710-637-3	Sequence 3, Appli
45	1514	86.7	7298	5	PCT-US93-00907-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-09-721-480-6  
; Sequence 6, Application US/09721480  
; Patent No. 6740323  
; GENERAL INFORMATION:  
; APPLICANT: Selby, Mark  
; APPLICANT: Glazer, Edward  
; TITLE OF INVENTION: HBV/HCV VIRUS-LIKE PARTICLE  
; FILE REFERENCE: PP01635.002  
; CURRENT APPLICATION NUMBER: US/09/721,480  
; CURRENT FILING DATE: 2000-11-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 5882  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: plasmid  
; OTHER INFORMATION: PCMV-II-E2661-sag  
; NAME/KEY: CDS  
; LOCATION: (1992)..(3584)  
US-09-721-480-6

Alignment Scores:  
Pred. No.: 1.07e-153  
Score: 1746.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
Matches: 303  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-09-721-480-6 (1-5882)

QY 1 MetAspAlaMetIysArgGlyLeuCysValLeuLeuLeuCysGlyAlaValPheVal 20

```

1992 ATGGATGCAATGAAGAGAGGCTCTGCTGTGTGCTGCTGCTGTGGAGCACTCTCGTT 2051
QY 21 SerProSerAlaSerGluThrHisValThrGlySerAlaGlyHisThrValSerGly 40
Db 2052 TCGCCAGCGCTAGCGAACCACACGTCACCGGGGAAGTCCCGGCCACACTGTGCTGGA 2111
QY 41 PheValSerLeuLeuAlaProGlyAlaLysGlnAenValGlnLeuIleAenThrAenGly 60
Db 2112 TTTGTTAGCTCTCTGCACAGCGCCCAAGCAGAGCTCCAGCTGATCAACACCAACGCG 2171
QY 61 SerTrpHisLeuAenSerThrAlaLeuAenCysAenAenSerLeuAenThrGlyTrpLeu 80
Db 2172 AGTTGGCAGCTCATAGCAGCGCCCTGAACCTGCAATGATAGCTCAACACCGCGCTGGT 2231
QY 81 AlaGlyLeuPheTyrHisHisPheAenSerSerGlyCysProGluArgLeuAlaSer 100
Db 2232 GCAGGGCTTTTCTATCACCAAGTTCAACTCTTCAGGCTGTCTTGAGAGGCTAGCCAGC 2291
QY 101 CysArgProLeuThrAspPheAspGlnGlyTrpGlyProIleSerTyrAlaAenGlySer 120
Db 2292 TCGCGACCCCTTACCGAATTTGACACAGGCTGGGGCCCTATCAGTTATGCCAACCGGAAC 2351
QY 121 GlyProAspGlnArgProTyrCysTrpHisTyrProProLysProCysGlyIleValPro 140
Db 2352 GGGCCCGACAGCGCCCTACTGTGCGACTACCCCGCCCAACCTTGCCTATTTGGCCC 2411
QY 141 AlaLysSerValCysGlyProValTyrCysPheThrProSerProValValGlyThr 160
Db 2412 GCGAAGAGTGTGTGCGGTATATTCTTCACTCCCGAGCCCGTGGTGGGAACG 2471
QY 161 ThrAspArgSerGlyAlaProThrTyrSerTrpGlyGluAenAspThrAspValPheVal 180
Db 2472 ACCGACAGTTCGGCGCGCCACCTACAGCTGGGGTGAAAAATGATACGGACGCTTTCGTC 2531
QY 181 LeuAenAenThrArgProProLeuGlyAenTrpPheGlyCysThrTrpMetAenSerThr 200
Db 2532 CTTAACAAATACAGGCCACCGTGGGCAATTTGTTCCGTTGTACCTGGATGAACCTCACT 2591
QY 201 GlyPheThrLysValCysGlyAlaProProCysValIleGlyAlaGlyAenAenThr 220
Db 2592 GGATTCACCAAGTGTGGAGCGCTCTCTTGTGTATCGGAGGGCGGCAACACACC 2651
QY 221 LeuHisCysProThrAspCysPheArgLysHisProAspAlaThrTyrSerArgCysGly 240
Db 2652 CTGCACTGCCCACTGATGCTTCCGCAAGCATCCGAGCGCCACATACCTCTCGGTGCGGC 2711
QY 241 SerGlyProTrpIleThrProArgCysLeuValAspTyrProTyrArgLeuTrpHisTyr 260
Db 2712 TCCGTCCTCTGGATCACCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2771
QY 261 ProCysThrIleAenTyrThrIlePheLysIleArgMetTyrValGlyGlyValGluHis 280
Db 2772 CCTTGTACCATCACTACACCATATTTAAATCAGGATGTACGTGGAGGGGTGCAACAC 2831
QY 281 ArgLeuGluAlaAlaCysAenTrpThrArgGlyGluArgCysAspLeuGluAspArgAsp 300
Db 2832 AGGCTGGAAGCTCCCTGCAACTGACCGGGGGGAACGTTGGCATCTGGAAGATAGGAC 2891
QY 301 ArgSerGlu 303
Db 2892 AGGTCCGAG 2900

```

RESULT 2

```

US-08-460-806-12
; Sequence 12, Application US/08460806
; Patent No. 5747241
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: HARADA, SHIZUKO
; APPLICANT: HONDA, YOSHIKAZU
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS C

```

```

; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,806
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,630
; FILING DATE: 19-OCT-1994
; APPLICATION NUMBER: US 07/956,993
; FILING DATE: 06-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5747241man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4667-001-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C virus
; IMMEDIATE SOURCE:
; CLONE: HCV1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1207
; US-08-460-806-12

```

```

Alignment Scores:
Pred. No.: 6,63e-143 Length: 1207
Score: 1620.50 Matches: 283
Percent Similarity: 96.9% Conservative: 1
Best Local Similarity: 96.5% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 1 Gaps: 1

```

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-08-460-806-12 (1-1207)

```

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 104 GTGCTGCTGCTATTTGCCGGCGTC-----GACGCGGAACCCACGTCACC 148
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 149 GGGGGAAGTGGCGGCCACACTGTGCTGGATTGTTAGCTCTCTGCACACGCGCCAAG 208
QY 51 GlnAenValGlnLeuIleAenThrAenGlySerTrpHisLeuAenSerThrAlaLeuAen 70
Db 209 CAGACGCTCAGCTGATCAACACCAACGCGAGTGGCAGCTCAATAGCACGCGCCCTGAAC 268
QY 71 CysAenAspSerLeuAenThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAen 90
Db 269 TGCATATGATAGCTCAACACCGCGCTGGTGGCAGGCGCTTTTCTATCACCACCAAGTTCAAC 328

```



```

QY 91 SerSerGlyCysProGluAargLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
DB 329 TCITCAGCGTCTCTGAGAGCTAGCCAGCTGCCGACCCCTTACCGATTTCAGCAGGC 388
QY 111 TrpGlyProIleSerTyraAlaGlySerGlyProAspGlnArgProTyrcysTrpHis 130
DB 389 TGGGGCCCTATCAGTTATGCCAACGGAAGGGCCCGCAGCAGGCCCTACTGCTGGCAC 448
QY 131 TyrProProIysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
DB 449 TACCCCAAAACCTTCGGTATTTCGCCCGGAGAGTGTGTGTCGGTATATTGC 508
QY 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
DB 509 TTCACCTCCAGCCCGTGGTGGGAACGACGACAGGTGGGGCGGCCCTACACG 568
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProLeuGlyAsn 190
DB 569 TGGGGTGAANAATGATACGAGCGTCTTCGTCCTTAACAATAACGAGCCGCTGGGCAAT 628
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
DB 629 TGGTTCGGTTGACCTGGATGAACCTCAACTGGATTACCAAGTGTGGGAGCGCTCCT 688
QY 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
DB 689 TGTGTCTATCGAGGGGGGGCAACACACCTGCACCTGCCCTGATTCCTCCGCAAG 748
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
DB 749 CATCCGAGCCACATACCTCGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 808
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnThrThrIlePheLys 270
DB 809 GTGACCTACCCGATAGGCTTTGGCATTTCTTGTACCATCACTACCATATTATAA 868
QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
DB 869 ATCAGGATGATCGTGGGGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 928
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
DB 929 GGGCAACGTTGCGATCTGGAAGACAGGACAGGTCCGAG 967

```

RESULT 3

```

US-08-325-630-12
; Sequence 12, Application US/08325630
; Patent No. 5750331
; GENERAL INFORMATION:
; APPLICANT: MIYAMURA, TATSUO
; APPLICANT: SAITO, IZUMU
; APPLICANT: HARADA, SHIZUKO
; APPLICANT: HONDA, YOSHIKAZU
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS C
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/325,630
; FILING DATE:
; CLASSIFICATION: 435

```

PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 07/956,993
; FILING DATE: 06-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5750331man P.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 4667-001-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C virus
; IMMEDIATE SOURCE:
; CLONE: HCV1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1207
; US-08-325-630-12

```

Alignment Scores:

```

Pred. No.: 6,63e-143 Length: 1207
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 1 Gaps: 1

```

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-08-325-630-12 (1-1207)

```

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
DB 104 GTGCTGCTGCTATTTCCCGCGTC-----GACGGGAAACCCACGTCACC 148
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
DB 149 GGGGGAAGTCCGCGCCACACTGTGTCTGGATTGTTAGCTCTCTCGCACCAGCGCCAA 208
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
DB 209 CAGAACGTCAGCTGATCAACACCAACGCGCAGTTGGCACCCTCAATAGCACGCGCCCTGAAC 268
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
DB 269 TGCATGATGAGCTCAACACCGCGCTGTTGGCAGGGCTTTTCTATCACCACAGTTCAAC 328
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
DB 329 TCTTCAGGCTGTCTCGAGAGGCTAGCAGCTGCGGACCCCTTACCGATTTCAGCAGGC 388
QY 111 TrpGlyProIleSerTyraAlaGlySerGlyProAspGlnArgProTyrcysTrpHis 130
DB 389 TGGGGCCCTATCAGTTATGCCAACGGAAGGGCCCGCAGCAGGCCCTACTGCTGGCAC 448
QY 131 TyrProProIysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
DB 449 TACCCCAAAACCTTCGGTATTTCGCCCGGAGAGTGTGTGTCGGTATATTGC 508
QY 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
DB 509 TTCACCTCCAGCCCGTGGTGGGAACGACGACAGGTGGGGCGGCCCTACACG 568
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProLeuGlyAsn 190
DB 569 TGGGGTGAANAATGATACGAGCGTCTTCGTCCTTAACAATAACGAGCCGCTGGGCAAT 628

```



```

RESULT 5
US-07-965-285-35
; Sequence 35, Application US/07965285
; Patent No. 5879904
; GENERAL INFORMATION:
; APPLICANT: Brecht, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF INVENTIONS: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; DESCRIPTION: cDNA to genomic RNA
US-07-965-285-35

Alignment Scores:
Pred. No.: 6.65e-143 Length: 1210
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservativeness: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 2 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-07-965-285-35 (1-1210)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
DB 122 GTGCTGCTGCTATTTGCCGGCGTC-----GACGCGGAACCCACGCTCACC 166
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
DB 167 GGGGGAAGTCCGGCCGACACTGTGTGATTGTAGCTCTCGACAGCGCCCAAG 226
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
DB 227 CAGAAGCTCAGCTGATCAACACCAACGCGCACTTGGCACTCAATAGCAGCGCTCTGAC 286
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheThrHisHisLysPheAsn 90

```

```

DB 287 TGCATGATAGCGCTTAACACCGCGCTGGTTGGCAGGCGCTTTCTATCACCACCAAGTTCAAC 346
QY 91 SerSerGlyCysProGluAlaSerCysArgProLeuThrAspPheAspGlnGly 110
DB 347 TCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCCGATTTTGACCAAGGC 406
QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
DB 407 TGGGGCCCTATCAGTTATGCCAACGGAAGGGCCCGACGCGCCCTACTGCTGGCAC 466
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
DB 467 TACCCCCCAAAACCTTGGCGTATTGTGCCCGGAGAGTGTGTGGTATATATTCG 526
QY 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
DB 527 TTCACCTCCAGCCCCGTGTGTGGAGACGACGAGTGGCGCGCGCCACCTACAGC 586
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
DB 587 TGGGGTGAATAATGATACGAGCGCTTCGTCTTAAACAATACCAAGGCCACCGCTGGGCAAT 646
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
DB 647 TGGTTGGGTTGACCTGGATGAACCACTGGATTTCACCAAAAGTGTGGGAGCGGCTCCT 706
QY 211 CysValIleGlyValAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
DB 707 TGTGTATCGAGGGGCGGCGCAACACCCCTGCACTGCCCTCCCTGATTTGCTTCCCAAG 766
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
DB 767 CATCCGGAGCCACATACCTCTCGGTGGCGCTCCGCTCCCTGGATCACACCCAGTGGCTG 826
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
DB 827 GTCGACTACCCGATATAGGCTTTGGCATTTATCCTTGTACCATCAACTACCATATTTAAA 886
QY 271 IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
DB 887 ATCAGATGTAGTGGAGGGGTGCAACACAGCTGGAAAGCTGCTTGCACCTGGACGCGG 946
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
DB 947 GCGCAACGTTGCGATCTGGAAGACAGGACAGCTCCGAG 985

RESULT 6
US-08-487-231-35
; Sequence 35, Application US/08487231
; Patent No. 5919454
; GENERAL INFORMATION:
; APPLICANT: Brecht, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF INVENTIONS: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,231

```

```

; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; DESCRIPTION: cdna to genomic RNA
; US-08-487-231-35

Alignment Scores:
Pred. No.: 6,65e-143 Length: 1210
Score: 1620.50 Matches: 283
Percent Similarity: 96.91% Conservative: 1
Best Local Similarity: 96.58% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 2 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-08-487-231-35 (1-1210)
QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
DB 122 GTGCTGCTCTATTGTCGGCGCTC-----GACGGGAACCCACGTCACC 166
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
DB 167 GGGGGAAGTCCGGCCACACTGTCTGGATTGTGTAGCTCTCGCACGAGCGCCAAG 226
QY 51 GlnAsnValcIcLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
DB 227 CAGAACGTCAGCTGATCAACCAACACGCGAGTTGGCACCTCAATAGCACGGCTCTGAAC 286
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyRHisLeuLysPheAsn 90
DB 287 TGCAATGATAGCCTTAACACCGGCTGGTTGGCAGGGCTTTCTATCACCACCAAGTTCAAC 346
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
DB 347 TCTTACAGGTCTCTCAGAGGCTAGCCAGCTGCCAGCCCTTACCGATTTCACAGGGGC 406
QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
DB 407 TGGGGCCCTATCAGTTATGCAACGGAAGCGGCCCGACGAGCGCCCTACTGCTGGCAC 466
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
DB 467 TACCCCCCAAAACCTTGGCGGTATTGTGCCCGCAAGAGTGTGTGGTCCGGTATATTGC 526
QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
DB 527 TTCACTCCCAAGCCCCGTGGTGGTGGGAACGACCGACAGGTCCGGCGGCCCACTACAGC 586
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
DB 587 TGGGGTGAATAATGATACGACGCTTCCTGCTCTTAAACAATACCGAGGCCCGCTGGGCAAT 646
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210

```

```

Db 647 TGGTTCCGTTGTACCTCGGATGAACCTCAACTGGATTCCAAAGTGTCCGAGCGCCCTCCT 706
QY 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 707 TGTGTCAATCGAGGGGGCGGCAACACACCTGTCACCTGCCCACTGATTGCTTCGCAAG 766
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 767 CATCCGACGCCACATACTCTCGGTGGGCTCCGGTCCCTGGATCACACCCAGGTGCTG 826
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 827 GTCCACTACCCGTTATAGCTTTGGCAATTATCTTTGTACCATCAACTACCATATTTAAA 886
QY 271 IleArgHetTyrValGlyGlyValGluHisArgLeuGluAlaCysAsnTrpThrArg 290
Db 887 ATCAGGATGTACGTGGGAGGGGTGAAACACAGGCTGGAAGCTGCTGCNACTGGACGGG 946
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 947 GCGGACGTTCCGATCTGGAGACAGGACAGGTCCGAG 985

RESULT 7
US-09-201-912-35
; Sequence 35, Application US/09201912
; Patent No. 6210962
; GENERAL INFORMATION:
; APPLICANT: Brecht, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,285
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
; DESCRIPTION: cdna to genomic RNA
; US-09-201-912-35

Alignment Scores:

```

```

Pred. No.: 6.65e-143 Length: 1210
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 3 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-09-201-912-35 (1-1210)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 122 GTCTGCTGCTATTTCGCGCGTC-----GACCGGAACCCACGTCACC 166

QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 167 GGGGGAAGTCCGCGCCACACTGTGTCTGGATTGTTAGCTCTCTGCACCGCGCCAG 226

QY 51 GlnAsnValGlnLeuLeuAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 227 CAGAAGCTCCAGCTGATCAACACCAACGGCAGTTGGCACCCTCAATAGCAGCGCTCTGAAC 286

QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrrHisIlysPheAsn 90
Db 287 TGCATGATAGCTTTAACCACCGCTGTTGGCAGGCTTTTCTATCACCACAGTTTCAAC 346

QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 347 TCTTCAGCTGCTCTCAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTGACACGAGGC 406

QY 111 TrpGlyProIleSerTyrrAlaAsnGlySerGlyProAspGlnArgProTyrrCysTrpHis 130
Db 407 TGGGGCCCTATCATGTTATGCAACGGAAGCGGCGCCGACGAGCGCCCTACTGCTGGCAC 466

QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrrCys 150
Db 467 TACCCCCCAAACTTGGGTATTGTCGCCGGAAGAGTGTGTGTCGGTATATTGC 526

QY 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrrSer 170
Db 527 TTCCTCCAGCCCGCTGTTGGTGGGAACGACGACGAGCTGGCGCGCCACCTACAGC 586

QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 587 TGGGGTGAATATGATACGAGACGCTTTCGTCTTAAACATACCGAGCCACCGCTGGGCAAT 646

QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 647 TGGTTCGGTTTACCTTGGATGAACCTCACTGGATTCCCAAGTGTGGGAGCGCTCTCT 706

QY 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 707 TGTGTCATCGGAGGGCGGGCAACACACCTGCACCTGCCACCTGATTGCTTCGCGAAG 766

QY 231 HisProAspAlaThrTyrrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 767 CATCCGAGCCACATCATCTCGGTGCGGCTCCGCTCCCTGGATCACACCGAGTGCCTG 826

QY 251 ValAspTyrrProTyrrArgLeuTrpHisTyrrProCysThrIleAsnTyrrThrIlePheLys 270
Db 827 GTGACATACCGGTATAGGCTTTGGCATATCTCTTGACCATCACTACACCATATTATAA 886

QY 271 IleArgMetTyrrValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
Db 887 ATCAGATGATAGTGGAGGGGTCCGACACAGCTGGAAGCTGCTCGCACTGGAGCGG 946

QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 947 GCGCAACGTTGGATCTGGAAGACAGGAGGAGGAGTCCGAG 985

```

RESULT 8

US-08-824-057-4

; Sequence 4, Application US/08824057

; Patent No. 6121020

```

; GENERAL INFORMATION:
; APPLICANT: SELBY, MARK
; APPLICANT: HOUGHTON, MICHAEL
; TITLE OF INVENTION: NOVEL HEPATITIS C E1 AND E2 TRUNCATED
; TITLE OF INVENTION: POLYPEPTIDES AND METHODS OF OBTAINING THE SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,057
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,959
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0987.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1989 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-824-057-4

```

```

Alignment Scores:
Pred. No.: 1.38e-142 Length: 1989
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 3 Gaps: 1

```

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-08-824-057-4 (1-1989)

```

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 31 GTGCTGCTGCTATTTCGCGCGTC-----GACCGGAACCCACGTCACC 75

QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 76 GGGGGAAGTCCGCGCCACACTGTGTCTGGATTGTTAGCTCTCTGCACCGCGCCAG 135

QY 51 GlnAsnValGlnLeuLeuAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 136 CAGAAGCTCCAGCTGATCAACACCAACGGCAGTTGGCACCCTCAATAGCAGCGCTGAAC 195

QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrrHisIlysPheAsn 90
Db 196 TGCATGATAGCTTTAACCACCGCTGTTGGCAGGCTTTTCTATCACCACAGTTTCAAC 255

QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 256 TCTTCAGCTGCTCTCAGAGGCTAGCCAGCTGCGAGCCCTTACCGATTGACACGAGGC 315

QY 111 TrpGlyProIleSerTyrrAlaAsnGlySerGlyProAspGlnArgProTyrrCysTrpHis 130

```

```

Db      316 TGGGGCCCTATCAGTTATGCCAAGAGCGCGCCGACCGCCCTACTGCTGGCAC 375
Qy      131 TyrProProLysProCysGlyValProAlaLysSerValCysGlyProValTyrCys 150
Db      376 TACCCCAAAACCTTGGCGTATTGTGCCCGGAGAGTGTGTGGTTCGGTATATTGC 435
Qy      151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db      436 TTCACCTCCAGCCCGTGGTGGGACGACCGACAGTGGGGCGCCACCTACAGC 495
Qy      171 TrpGlyGluAsnAspThrAspValPheValLeuLeuAsnThrArgProProLeuGlyAsn 190
Db      496 TGGGGTGAAATGATACGACGCTCTCGTCTTAACATACAGCCACCGCTGGCAAT 555
Qy      191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db      556 TGGTTCGGTTGTACCTGGATGAACCAACTGGATTACCAAAAGTGTGGGAGCGCTCCT 615
Qy      211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db      616 TGTGTATCGGAGGCGCGGCAACACACCTGCACTCCGCTCCGTCACACCGAGTTCGCAAG 675
Qy      231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db      676 CATCCGGAGCCACATATCTCGTTCGGCTCCGCTCCGTCACACCGAGTTCGCTG 735
Qy      251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db      736 GTCCGACTACCCGCTATAGGCTTTGGCATTTATCTTGTACCATCAACTACATATATTAA 795
Qy      271 IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
Db      796 ATCAGGATGTACGTGGGAGGGGTGCAGACAGCTGGAGCTGCCTGCACCTGGACGCGG 855

Qy      291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db      856 GCGCAACGTTGCACTGGAAGATAGGACAGTCCGAG 894

RESULT 9
US-09-415-582-4
; Sequence 4, Application US/09415582
; Patent No. 6326171
; GENERAL INFORMATION:
; APPLICANT: SELBY, MARK
; HOUGHTON, MICHAEL
; TITLE OF INVENTION: NOVEL HEPATITIS C E1 AND E2 TRUNCATED
; POLYPEPTIDES AND METHODS OF OBTAINING THE SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/415,582
; FILING DATE: 08-Oct-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0987,006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

```

```

;
; LENGTH: 1989 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-415-582-4

Alignment Scores:
Pred. No.: 1,38e-142 Length: 1989
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
Gaps: 3

US-10-715-665-7_COPY_1_303 (1-303) x US-09-415-582-4 (1-1989)

Qy      11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db      31 GTGCTGCTGCTATTGCGCGGCTC-----GACGGCGAAACCCACGCTCAC 75
Qy      31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyValAlaLys 50
Db      76 GGGGGAAGTGGCGGCCACACTGTGTGTGATTTGTTAGCCTCCTCGCACCGAGCGCCAAAG 135
Qy      51 GluAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db      136 CAGAACGTCCAGCTGATCAACCAACCGCAGTTGGCACCTCAATAGCACGGCCCTGAAC 195
Qy      71 CysAsnAspSerLeuLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisIleAspPheAsn 90
Db      196 TGCAATGATAGCTCAACACCGCGCTGGTGGCAGGGCTTTCTATCATCACCAAGTTCAAC 255
Qy      91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db      256 TCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCGGACCCCTTACCGATTTTGACCGAGGC 315
Qy      111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db      316 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCGCCCGACCGCCCTACTGCTGGCAC 375
Qy      131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db      376 TACCCCAAAACCTTGGCGTATTGTGCCCGAAGAGTGTGTGGTCCGGTATATTGC 435
Qy      151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db      436 TTCACCTCCAGCCCGTGGTGGGAAACGACCGACAGTGGGGCGCCACCTACAGC 495
Qy      171 TrpGlyGluAsnAspThrAspValPheValLeuLeuAsnThrArgProProLeuGlyAsn 190
Db      496 TGGGGTGAAATGATACGACGCTCTCGTCTTAACATACACCGACCGCCGCTGGCAAT 555
Qy      191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db      556 TGGTTCGGTTGTACCTGGATGAACCAACTGGATTACCAAAAGTGTGGGAGCGCTCCT 615
Qy      211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db      616 TGTGTATCGGAGGCGCGGCAACACACCTGCACTCCGCTCCGTCACACCGAGTTCGCAAG 675
Qy      231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db      676 CATCCGGAGCCACATATCTCGTTCGGCTCCGCTCCGTCACACCGAGTTCGCTG 735
Qy      251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db      736 GTCCGACTACCCGCTATAGGCTTTGGCATTTATCTTGTACCATCAACTACATATATTAA 795
Qy      271 IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
Db      796 ATCAGGATGTACGTGGGAGGGGTGCAGACAGCTGGAGCTGCCTGCACCTGGACGCGG 855

```

Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
Db 856 GCGGAACGTTGGCATCTGGAAGATAGGACACAGGTCGGAG 894

## RESULT 10

US-09-693-596-3  
; Sequence 3, Application US/09693596  
; Patent No. 6521423  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Abrignani, Sergio  
; APPLICANT: Chien, David  
; APPLICANT: Selby, Mark  
; APPLICANT: Glazer, Edward  
; TITLE OF INVENTION: Intracellular Production of Hepatitis C E1 and E2  
; FILE REFERENCE: 1378.002  
; CURRENT APPLICATION NUMBER: US/09/693,596  
; CURRENT FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/073,406  
; PRIOR FILING DATE: 1998-05-06  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1989  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1989)  
US-09-693-596-3

## Alignment Scores:

Pred. No.: 1.38e-142 Length: 1989  
Score: 1620.50 Matches: 283  
Percent Similarity: 96.93% Conservative: 1  
Best Local Similarity: 96.59% Mismatches: 4  
Query Match: 92.81% Indels: 5  
DB: 4 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-09-693-596-3 (1-1989)

Qy 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
Db 31 GTGCTGCTGCTATTGCGCGCTC-----GACGGGAAACCCAGTCACC 75  
Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
Db 76 GGGGGAAGTCCGGCCACACTGTGCTGGATTGTTAGCTCTCGCACCGGCGCAAG 135  
Qy 51 GluAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70  
Db 136 CAGAACGTCAGCTGATCAACACACCGGAGTGGCACCTCAATAGCAGCGCCCTGAAC 195  
Qy 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheThrHisHisLysPheAsn 90  
Db 196 TGCATGATAGCTTCAACACCGGCTGGTGGAGGGCTTTCTATCACCAAGTTCAAC 255  
Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
Db 256 TCCTCAGGCTGCTCTGAGAGGCTAGCCAGCTGCGGACCCCTTACCGATTGTTGACCGGCG 315  
Qy 111 TrpGlyProIleSerThrAlaAsnGlySerGlyProAspGlnArgProThrCysTrpHis 130  
Db 316 TGGGGCCCTATCAGTTATGCAACGGAAGCGGCGCCGACGAGCGCCCTACTCTGCGCAC 375  
Qy 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValThrCys 150  
Db 376 TACCCCAAAACCTTGGGATATGTGCCCGGAGAGTGTGTGTCGGGTATATTC 435  
Qy 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrThrSer 170

Db 436 TTCACCTCCAGCCCGCTGGTGGTGGAAACGACACAGGTCGGGCGGCCACCTACAGC 495  
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProLeuGlyAsn 190  
Db 496 TGGGGTGAAATGATACGGACGCTTCGTCTTAAACAATACCGGCCACCGCTGGCAAT 555  
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
Db 556 TGGTTCGGTTGTACCTGGATGAACCTCACTGGATTACCAAGTGTGGGAGCGCCCTCCT 615  
Qy 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
Db 616 TGTGTGTCATCGGAGGGCGGCAACACACCTGCACTGCGCCCACTGATTGTTCCGCAAG 675  
Qy 231 HisProAspAlaThrTrpSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
Db 676 CATCCGAGCCACATACTCTCGTGGCGCTCCGGTCCCTGGATCACACCCAGGTGCTG 735  
Qy 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
Db 736 GTCGACTACCCGTATAGGCTTTGGCATTATCTTGTACCATCACTACACCATATTTAA 795  
Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaLysCysAsnTrpThrArg 290  
Db 796 ATCAGGATGTACGTGGAGGGGTGGAACACAGGCTGGAAGCTGCTGCAACTGGACGCG 855  
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
Db 856 GCGGAACGTTGGCATCTGGAAGATAGGACACAGGTCGGAG 894

## RESULT 11

US-08-444-818-88  
; Sequence 88, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 88:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8316 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA



```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8316
US-08-444-818-88

Alignment Scores:
Pred. No.: 1,14e-141      Length: 8316
Score: 1620.50           Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81%      Indels: 5
DB: 3                    Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-08-444-818-88 (1-8316)
QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 778 GTGCTGCTGCTATTTCGGCGGCTC-----GACGCGGAAACCCACGTCACC 822
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 823 GGGGGAAGTCCGCGGCACACTGTGTCTGGATTGTGTAGCTCTCTCGCACCGCGCCAAAG 882
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 893 CAGAAGCTCAGCTGATCAACCAACGAGGTTGGCACCTCAATAGCACGGCCCTGAAC 942
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
Db 943 TGCATGATAGCTCAACACCGGCTGTTGGCAGGCTTTTCTATCACCAAGTTCAAC 1002
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1003 TCTTCAGGCTGCTCTGAGAGGCTAGCCAGCTGCGACCCCTTACCGATTTCACCAAGGCG 1062
QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 1063 TGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCCGACAGCGGCCCTACTGCTGGCAC 1122
QY 131 TyrProPolysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 1123 TACCCCCAAAACCTTTGCGGTATTGTGCGCGGAGAGTGTGTGGTCCGGTATATTGC 1182
QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1183 TTCACCTCCAGCCCGGTGTGTGGGAAACGACGAGTGGCGGCGCCCACTACAGC 1242
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1243 TGGGGTGMAAATGATACGGACGCTTCGTCTTAACATACAGGCGCACCGCTGGGCAAT 1302
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1303 TGGTTCCGTTGTACCTGGATGAACCTCAACTGGATTTCACCAAGTGTGCGAGCGCCTCCT 1362
QY 211 CysValIleGlyAlaGlyAlaAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 1363 TGTGTCTATCGAGGGCGGGCAACACACCTGCCTGCCCCCACTGATTGCTTCCGCAAG 1422
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 1423 CATCCGACGCCACATACTCTCGGTGGGCTCCGGTCCCTGGATCAACCCAGTGCCTTG 1482
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 1483 GTCCACTACCCGATAGGCTTTGSCATTATCTTTGTACCATCAACTACCATATTTAAA 1542
QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaCysAsnTrpThrArg 290
Db 1543 ATCAGGATGTACGTGGAGGGGTGGAACACAGAGCTTGAAGCTCCCTGCAACTGGACGCG 1602
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
```

```
Db 1603 GCGGAACGTTCGATCTGGAAGACAGGACAGGTCGAG 1641

RESULT 12
US-08-444-818-137
; Sequence 137, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANEV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 137:
; LENGTH: 8987 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8985
; US-08-444-818-137

Alignment Scores:
Pred. No.: 1,28e-141      Length: 8987
Score: 1620.50           Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81%      Indels: 5
DB: 3                    Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-08-444-818-137 (1-8987)
QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1120 GTGCTGCTGCTATTTCGGCGGCTC-----GACGCGGAAACCCACGTCACC 1164
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1165 GGGGGAAGTCCGCGGCACACTGTGTCTGGATTGTGTAGCTCTCTCGCACCGCGCCAAAG 1224
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 1225 CAGAAGCTCAGCTGATCAACCAACGAGGTCGACCTCAATAGCACGGCCCTGAAC 1284
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
```

```
Db 1285 TGCAATGATAGCCTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACCAAGTTCAAC 1344
Qy 91 SerSerGlyCysProGluAuaGLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1345 TCITTCAGGCTGCTCTGAGAGCTAGCAGCTGCCGACCCCTTACCGATTTCACAGGGC 1404
Qy 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 1405 TGGGGCCCTATCAGTTATGCAACCGAAGCGGCCCGCCAGCCAGCGCCCTACTGCTGGCAC 1464
Qy 131 TyrProProlysProCysGlyValProAlaLysSerValCysGlyProValTyrCys 150
Db 1465 TACCCCAAAACCTTCGGGTATTGTGCCCGGAAGAGTGTGTGGTCCGGTATATTC 1524
Qy 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1525 TTCACTCCAGCCCCGTTGGTGGGAACCGACGAGGTCCGGCGGCCCCACTACAGC 1584
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1585 TGGGGTGAAAATGATACGGACGCTTCCTCCTTAACAATACAGGCCACCGCTGGGCAAT 1644
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1645 TGGTTCGGTTGTACCTGGATGAACCTCAACTGGGATTCCAAAGGTGCGGAGCGCCTCT 1704
Qy 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 1705 TGTGTATCGAGGGGGGGGCAACACCTGCACTGCCCCACTGATTCCTTCCGCGAAG 1764
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 1765 CATCCGAGCCACATACCTCGGTGGGCTCCGGTCCCTGGATCACACCAGGTGCTG 1824
Qy 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 1825 GTGCACTACCGGTATAGGCTTGGCATTTGTCATTTGTTGTTGTTGTTGTTGTTGTTG 1884
Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
Db 1885 ATCAGATGTACGTGGAGGGGTGGAACAGAGCTGGAACTGCTGCAACTGCAAGCGG 1944
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 1945 GCGCAAGCTTGGCATCTGGAAGACAGGACAGGACAGTCCGAG 1983
```

## RESULT 13

```
US-08-444-818-122
; Sequence 122, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Ruter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
```

```
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-444-818-122

Alignment Scores:
Pred. No.: 1,32e-141 Length: 9185
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 3 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-08-444-818-122 (1-9185)

Qy 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1439 GTGCTCTGCTATTTTCCCGCGGTC-----GACGCGGAAACCCACGTCAAC 1483
Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1484 GGGGGAGGTGCCCGCCACACTGTGTCTGGATTGTAGCTCTCTCGACACGCGCCAG 1543
Qy 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 1544 CAGAACGTCAGCTGATCAACACCAACGCGAGTTGGCACCTCAATAGCAGCGCCCTGAAC 1603
Qy 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
Db 1604 TGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCACCAAGTTCAAC 1663
Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1664 TCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCCCTTACCGATTTCACAGGGC 1723
Qy 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 1724 TGGGGCCCTATCAGTTATGCAACGAGCGGCCCGCCAGCCAGCGCCCTACTGCTGGCAC 1783
Qy 131 TyrProProlysProCysGlyValProAlaLysSerValCysGlyProValTyrCys 150
Db 1784 TACCCCAAAACCTTTCGGGTATTGTGCCCGGAAGAGTGTGTGGTCCGGTATATTC 1843
Qy 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1844 TTCACCTCCAGCCCCGTTGGTGGGAACGACGAGGTGGGGCGGCCCCACTACAGC 1903
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1904 TGGGGTGAAAATGATACGAGACGCTTCCTCGTCTTAAACAATACCGAGCCACCGCTGGGCAAT 1963
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1964 TGGTTCGGTTGTACCTGGATGAACCTCACTGGATTCCACCAAGTGTGGGAGCGCCTCT 2023
Qy 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 2024 TGTGTATCGAGGGGGGGGCAACACCTTGCACCTGCCACCTGATTTGCTTCCGCAAG 2083
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
```

Db 2084 CATCCGAGCCACATACCTCTCGGTCCGCTCGGTCCCTGGATCACACCCAGGTGCCTG 2143  
Qy 251 ValAspTyrProTyrArgLeuThrHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
Db 2144 GTGACTACCGGTATAGGCTTTGGCAATTATCTTGTTACCATCAACTACACCATATTTAAA 2203  
Qy 271 IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAsnTyrThrArg 290  
Db 2204 ATCAGGATGTAGCTGGAGGGGTGGAACACAGGCTGGAAGCTGCCTGCAACTGGACGGG 2263  
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
Db 2264 GCGCAACGTTCGATCTGGAACACAGGACAGGTCCGAG 2302

## RESULT 14

US-08-444-818-123/c  
; Sequence 123, Application US/08444818  
; Patent No. 6150087

## GENERAL INFORMATION:

; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777

## CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA

## COUNTRY: USA

; ZIP: 94608-2916

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/444,818

## FILING DATE:

; CLASSIFICATION: 424

## PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/403,590

## FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Harbin, Alisa A.

; REGISTRATION NUMBER: 33,895

; REFERENCE/DOCKET NUMBER: 0110.002

## TELECOMMUNICATION INFORMATION:

; TELEPHONE: (508)359-3876

; TELEFAX: (508)359-3885

; INFORMATION FOR SEQ ID NO: 123:

## SEQUENCE CHARACTERISTICS:

; LENGTH: 9185 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; ANTI-SENSE: YES

US-08-444-818-123

## Alignment Scores:

Pred. No.:	1-32e-141	Length:	9185
Score:	1620.50	Matches:	283
Percent Similarity:	96.93%	Conservative:	1
Best Local Similarity:	96.59%	Mismatches:	4
Query Match:	92.81%	Indels:	5
DB:	3	Gaps:	1

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-08-444-818-123 (1-9185)

Qy 11 ValLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
Db 7747 GTGCTGCTGCTATTTCGCGCGTC-----GACGCGGAACCCACGTCACC 7703

Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
Db 7702 GGGGGAAGTGGCGGCCACACATGTCTGGATTTGTAGCTTCTCTCGCACAGGCCCAAG 7643  
Qy 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70  
Db 7642 CAGAACGTCACGCTGATCAACACCAACGCGAGTTGGCACCTCAATAGCACGCGCCCTGAAC 7583  
Qy 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90  
Db 7582 TGCATATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATATCACACCAAGTTCAC 7523  
Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
Db 7522 TCTTCAGGCTGTCTGAGAGGCTAGCAGCTGCCGACCCCTTACCGATTTTGACCAAGGCG 7463  
Qy 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130  
Db 7462 TGGGCGCCCTATCAGTTATGCCAAGCGAGCGGCCCGACAGCGGCCCTACTGCTGGCAC 7403  
Qy 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
Db 7402 TACCCCCCAAAACCTTGGCGTATTGTGCCCGAAGAGTGTGTGGTCCGGTATATTGC 7343  
Qy 151 PheThrProSerProValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170  
Db 7342 TTCACCTCCAGCCCCGCTGGTGGGAACGACCGACAGGTGGGGCGGCCACCTACAGC 7283  
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
Db 7282 TGGGGTGAAATGATACGGAGCTCTTCGTCTCTTAACAATACCAAGCCACCGCTGGGCAAT 7223  
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
Db 7222 TGGTTCCGTTGTACCTGGATGAACCTCAACTGGATTACACAAAGTGTGGGAGCGGCTCT 7163  
Qy 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
Db 7162 TGTGTCATCGGAGGGCGGGGCAACACACCTGCACTGCCCCACTGATTGCTTCGCAAG 7103  
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
Db 7102 CATCCGAGCGCCACATACCTCTCGGTGCGGCTCCGTCCTGGATCACACCCAGGTGCTG 7043  
Qy 251 ValAspTyrProTyrArgLeuThrHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
Db 7042 GTTGACTACCCGCTATAGGCTTTGGCAATTATCTTGTACCATCAACTACACCATATTTAA 6983  
Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
Db 6982 ATCAGGATGTAGCTGGAGGGGTGGAACACAGGCTGGAAGCTGCCTGCAACTGGACGCG 6923  
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
Db 6922 GCGCAACGTTCGATCTGGAACACAGGACAGGTCCGAG 6884

## RESULT 15

US-09-388-874-1

; Sequence 1, Application US/09388874

; Patent No. 6284249

; GENERAL INFORMATION:

; APPLICANT: Veronique Barban  
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR

; TITLE OF INVENTION: TREATING C HEPATITIS

; FILE REFERENCE: PMCF97-03A

; CURRENT APPLICATION NUMBER: US/09/388,874

; EARLIER FILING DATE: 1999-09-02

; EARLIER APPLICATION NUMBER: PCT/FR98/00448

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 97/02,887

; EARLIER FILING DATE: 1997-03-06

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1  
; LENGTH: 9379  
; TYPE: DNA  
; ORGANISM: Virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (320)...(9352)  
US-09-388-874-1

## Alignment Scores:

Pred. No.: 1,36e-141 Length: 9379  
Score: 1620.50 Matches: 283  
Percent Similarity: 96.93% Conservative: 1  
Best Local Similarity: 96.59% Mismatches: 4  
Query Match: 92.81% Indels: 5  
DB: 3 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-09-388-874-1 (1-9379)

```
Qy 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1439 GTGCTGCTGCTATTGCGGGTC-----GACGCGGAACCCACGTCACC 1483
Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1484 GGGGGAAGTGGCGGCCACACTGTCTGGATTGTTAGCCTCTCTCGCACGAGGCGCAAG 1543
Qy 51 GluAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 1544 CAGAACGTCCAGCTGATCAACACCAACGCGAGTTGGCACCTCAATAGCAGGCGCCCTGAAC 1603
Qy 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90
Db 1604 TGCAATGATAGCTCAACACGGCTGGTGGCAGGGCTTTTCTATCACCAAGTTCAC 1663
Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1664 TCTTCAGGCTGTCTGAGAGCTAGCCAGCTGCCGCCCTTACCAGATTTTGACCAAGGC 1723
Qy 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 1724 TGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACGAGCGCCCTACTGCTGGCAC 1783
Qy 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 1784 TACCCCAAAACCTTCGGGTATTGTGCCCGGAGAGTGTGTGTCGGTATATTGC 1843
Qy 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1844 TTCACCTCCAGCCCGTGGTGGTGGGAACGACGACAGGTTCGGGCGGCCACCTACAGC 1903
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1904 TGGGTGAATAATGATACGACGCTTCGTCTTAACTAACATACGAGCCCGCTGGGCAAT 1963
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1964 TGGTTCGGTTGTACCTGATGAACCTCAACTGGATTACCAAGTGTGGGAGCGCTCCT 2023
Qy 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 2024 TGTGTATATCGGAGGCGGGCAACACACCTGTGCCTGCCCTCCCTGATTCCTCCGCAAG 2083
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 2084 CATCCGACGCCACATACTCTCGGTGGGCTCCGGTCCCTGGATCAACCCAGGTGCTG 2143
Qy 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 2144 GTGACTACCGGTATAGGCTTTGGCATTATCTTGTACCATCACTACACCATATTATA 2203
Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuAlaCysAsnTrpThrArg 290
```

Db 2204 ATCAGGATGTACGTGGGAGGGTGAACACACAGGCTGGAAGCTGCTGCACTGGACGGG 2263  
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
Db 2264 GCGAACGTTGCGATCTGGAAGACAGGGACAGGTCCGAG 2302

Search completed: April 15, 2005, 09:09:35  
Job time : 215 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: April 15, 2005, 08:04:43 ; Search time 1564 Seconds  
(without alignments)  
1175.328 Million cell updates/sec  
Title: US-10-715-665-7\_COPY\_1\_303  
Perfect score: 1746  
Sequence: 1 MDAMKRLCCVLLCGAVFV.....AACNWTGRGCDLDRDRSE 303  
Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 5622541 seqs, 303335566 residues  
Total number of hits satisfying chosen parameters: 11245082  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10715665/runat\_13042005\_165825\_28262/app\_query.fasta\_1.455  
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -LOCAL=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5  
-FGAPO=6 -FGAPEXT=7 -YGAPO=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7  
Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq\*  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq\*  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq\*  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1620.5	92.8	1914	15	US-10-187-257-3 Sequence 3, Appli
2	1620.5	92.8	1914	16	US-10-265-083-1 Sequence 1, Appli
3	1620.5	92.8	1989	17	US-10-371-040-3 Sequence 3, Appli
4	1620.5	92.8	9379	9	US-09-916-359-1 Sequence 1, Appli
5	1620.5	92.8	9401	18	US-10-445-724-1 Sequence 1, Appli
6	1523.5	87.3	9646	9	US-09-742-659-3 Sequence 3, Appli
7	1521.5	87.1	9416	9	US-09-438-076-19 Sequence 19, Appl
8	1521.5	87.1	9416	10	US-09-995-937-19 Sequence 19, Appl
9	1521.5	87.1	9416	10	US-09-917-563-19 Sequence 19, Appl
10	1521.5	87.0	9622	18	US-10-475-988-2 Sequence 2, Appli
11	1518.5	87.0	9365	10	US-09-827-688-7 Sequence 7, Appli
12	1515.5	86.8	9646	9	US-09-238-076-1 Sequence 1, Appli
13	1515.5	86.8	9646	10	US-09-995-937-1 Sequence 1, Appli
14	1515.5	86.8	9646	10	US-09-917-563-1 Sequence 1, Appli
15	1515.5	86.8	12980	9	US-09-238-076-5 Sequence 5, Appli
16	1515.5	86.8	12980	10	US-09-995-937-5 Sequence 5, Appli
17	1515.5	86.8	12980	10	US-09-917-563-5 Sequence 5, Appli
18	1512.5	86.6	1845	17	US-10-365-620-73 Sequence 73, Appl
19	1512.5	86.6	1845	19	US-10-912-968-75 Sequence 75, Appl
20	1512.5	86.6	2517	17	US-10-365-620-75 Sequence 75, Appl
21	1512.5	86.6	2517	19	US-10-912-969-77 Sequence 77, Appl
22	1512.5	86.6	2517	19	US-10-913-171-48 Sequence 48, Appl
23	1512.5	86.6	9599	17	US-10-189-359-13 Sequence 13, Appl
24	1512.5	86.6	10803	9	US-09-747-419-17 Sequence 17, Appl
25	1512.5	86.6	10803	15	US-10-259-275-17 Sequence 17, Appl
26	1503	86.1	1266	19	US-10-912-969-71 Sequence 71, Appl
27	1503	86.1	1938	17	US-10-365-620-71 Sequence 71, Appl
28	1503	86.1	1938	19	US-10-913-969-73 Sequence 73, Appl
29	1492.5	85.5	9416	9	US-10-913-171-46 Sequence 46, Appl
30	1492.5	85.5	9416	9	US-09-929-955-13 Sequence 13, Appl
31	1492.5	85.5	9416	13	US-10-104-966-13 Sequence 13, Appl
32	1492.5	85.5	9416	17	US-10-719-619-13 Sequence 13, Appl
33	1492.5	85.5	9416	18	US-10-817-591-13 Sequence 13, Appl
34	1398	80.1	750	9	US-09-952-572-1 Sequence 1, Appli
35	1393	79.8	834	19	US-10-488-269A-4 Sequence 4, Appli
36	1371.5	78.6	1266	17	US-10-365-620-69 Sequence 69, Appl
37	1371	78.5	834	19	US-10-488-269A-5 Sequence 5, Appli
38	1371	78.5	834	19	US-10-488-269A-6 Sequence 6, Appli
39	1324.5	75.9	9275	15	US-10-259-275-39 Sequence 39, Appl
C 40	1317	75.4	7927	15	US-10-128-590-32 Sequence 32, Appl
C 41	1317	75.4	7927	16	US-10-128-587A-32 Sequence 32, Appl
C 42	1317	75.4	7927	17	US-10-128-588B-32 Sequence 32, Appl
43	1313.5	75.2	3730	15	US-10-128-590-22 Sequence 22, Appl
44	1313.5	75.2	3730	16	US-10-128-587A-22 Sequence 22, Appl
45	1313.5	75.2	3730	17	US-10-128-578B-22 Sequence 22, Appl

ALIGNMENTS

RESULT 1  
US-10-187-257-3  
; Sequence 3, Application US/10187257  
; Publication No. US220030138458A1  
; GENERAL INFORMATION:  
; APPLICANT: HOUGHTON, Michael  
; APPLICANT: COATES, Steve  
; APPLICANT: O'HAGAN, Derek  
; TITLE OF INVENTION: HCV E1E2 VACCINE COMPOSITIONS  
; FILE REFERENCE: 2302-17206  
; CURRENT APPLICATION NUMBER: US/10/187,257  
; CURRENT FILING DATE: 2002-06-28  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1914  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: HCV-1 E1/E2/p7 region

889	Db	TGGGGCCCTACGTATTGCGCAACGGGAAGGGGCCCGACGAGCCCTACTGCTGGCAC	948
131	Qy	TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys	150
949	Db	TACCCCCCAAAACCTTGGGTATTGTGCCCGGGAAGAGTGTGTGGTCCGTATATTC	1008
151	Qy	PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer	170
1009	Db	TTCACTCCAGCCCGGTGGTGGGAACGACCACAGGTGGCGGGCCGACCTACAGC	1068
171	Qy	TrpGlyGluAsnAspThrAspValPheValIleuAsnAsnThrArgProProLeuGlyAsn	190
1069	Db	TGGGGTGAAATGATACGAGCGTCTTCGTCTTAAACAATACGAGGCACCGCTGGGCAAT	1128
191	Qy	TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro	210
1129	Db	TGGTTCGGTTGTACTCGATGAACCTCAACTGGATTACCAAAAGTGTGGGAGCGCCTCCT	1188
211	Qy	CysValIleGlyAlaLaglyAsnAsnThrIleuHisCysProThrAspCysPheArgLys	230
1189	Db	TGTGTCACTCGAGGGGGGGCAACAACCCCTGCACCTGCCCACTGATTGCTTCCGCAAG	1248
231	Qy	HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu	250
1249	Db	CATCCGGACGCCACATACCTTCGGTTCGGCTCCGGTCCCTGGATCAACACCGAGTGGCTG	1308
251	Qy	ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys	270
1309	Db	GTGCACTACCGGTATAGCTTTGGCAATTATCCTTGTACCACTCAACTACATATATTTAAA	1368
271	Qy	IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg	290
1369	Db	ATCAGGATGTACGTGGGAGGGGTGGAGCACAGGCTGGAAGCTGCCTGCACCTGACCGGG	1428
291	Qy	GlyGluArgCysAspLeuGluAspArgAspArgSerGlu	303



191 TtpPheGlyCysThrTtpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
 1129 TGGTTCGGTGTGTACCTGGATGAACCTCAACTGGATTCAACAAAGTGTGGGAGCGCTCT 1188  
 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 1189 TGTGTATCATCGAGGGGGGCAACACACCTGCACTGCCCTCACTGATTTCTCCGCAAG 1248  
 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTtpIleThrProArgCysLeu 250  
 1249 CATCCGAGCCACATACCTCTCGGTGGGTCCGTCCGTCCGTCCGTCCGTCCGTCCGT 1308  
 251 ValAspTyrProTyrArgLeuTtpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
 1309 GTGACTACCGGTATAGCTTTGGCATTAATCTTGTACCATCACTACATATATTAAA 1368  
 271 IleArgMetTyrValGlyGlyValGlyHisArgLeuGluAlaLysCysAsnTtpThrArg 290  
 1369 ATCAGGATGTACGTGGGAGGGGTGAGCAGCGTGGAGCTGGCACTGCCTGCAACTGGAGCGG 1428  
 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
 1429 GCGGACGTTGCCATCTGGAGATAGGACAGGTCCGAG 1467

RESULT 3

US-10-371-040-3  
 ; Sequence 3, Application US/10371040  
 ; Publication No. US2004001854A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Houghton, Michael  
 ; APPLICANT: Choo, Qui-Lim  
 ; APPLICANT: Abrignani, Sergio  
 ; APPLICANT: Chien, David  
 ; APPLICANT: Selby, Mark  
 ; APPLICANT: Glazer, Edward  
 ; TITLE OF INVENTION: Intracellular Production of Hepatitis C E1 and E2  
 ; FILE REFERENCE: 1378.002  
 ; CURRENT APPLICATION NUMBER: US/10/371,040  
 ; CURRENT FILING DATE: 2003-02-18  
 ; PRIOR APPLICATION NUMBER: US/09/073,406  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/045,675  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-06  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 1989  
 ; TYPE: DNA  
 ; ORGANISM: Hepatitis C virus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(1989)  
 ; US-10-371-040-3

Alignment Scores:  
 Pred. No.: 8,11e-167 Length: 1989  
 Score: 1620.50 Matches: 283  
 Percent Similarity: 96.93% Conservative: 1  
 Best Local Similarity: 96.59% Mismatches: 4  
 Query Match: 92.81% Indels: 5  
 DB: 17 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-10-371-040-3 (1-1989)

11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
 31 GTGCTGCTGCTATTTCGGCGCTC-----GACGCGAAACCCACGTCACC 75  
 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
 76 GGGGGAAGTCCCGGCCACACTGTGTGGATTGTGGCTCTCGCACCGCGCCAAAG 135

51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTtpHisLeuAsnSerThrAlaLeuAsn 70  
 136 CAGAACGTCAGCTGATCAACACCAACGCGAGTTGGCACCTCAATAGCAGCGCCCTGAAC 195  
 71 CysAsnAspSerLeuAsnThrGlyTtpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90  
 196 TGCATATGATAGCTCAACACCGCTGTGGTGGAGGGCTTTTCTATATCAACCAAGTTCAAC 255  
 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
 256 TCTTTCAGCTGTCTCGAGAGGCTAGCAGCTGCCGACCCCTTACCGATTTCACGAGGC 315  
 111 TtpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTtpHis 130  
 316 TGGGGCCCTATCATAGTTATGCCAAGCAAGCGCGCCCGACCGCCCTACTGCTGGCAC 375  
 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
 376 TACCCCCCAAAACCTTTGCGGTATTGTGCCCGCGAAGAGTGTGTGTCGGTATATTGC 435  
 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170  
 436 TTCACCTCCAGCCCCCGTGTGTGGGAACGACGACAGGTGGGGCGGCCACCTACAGC 495  
 171 TtpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
 496 TGGGGTGAAAAATGATACGGAGAGCTTTCGTCCTTAACAATACACGACCGCCCTGGGCAAT 555  
 191 TtpPheGlyCysThrTtpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
 556 TGGTTCGGTGTACCTGGATGAACCTCACTGGATTTCACCAAGTGTGGGAGCGCTCTCT 615  
 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 616 TGTGTCTATCGAGGGGGGGCAACAACACCTGCATCTGCCCACTGATTCCTCCGCAAG 675  
 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTtpIleThrProArgCysLeu 250  
 676 CATCCGAGCGCCACATACCTCTCGGTGGGTCCGCTCCGTGGATCAACCCAGGTGCTG 735  
 251 ValAspTyrProTyrArgLeuTtpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
 736 GTGACTACCGCTATAGGCTTTGGCATTAATCTTGTACCATCACTACACCATATTAAA 795  
 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaLysCysAsnTtpThrArg 290  
 796 ATCAGGATGTACGTGGAGGGGTGGAACACAGCTGGAAGCTGCTGCAACTGCAACGCGG 855  
 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
 856 GCGGACGTTGCCATCTGGAGATAGGACAGGTCCGAG 894

RESULT 4

US-09-916-359-1  
 ; Sequence 1, Application US/09916359  
 ; Patent No. US2002003473A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Veronique Barban  
 ; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR  
 ; TITLE OF INVENTION: TREATING C HEPATITIS  
 ; FILE REFERENCE: PMCF97-03A  
 ; CURRENT APPLICATION NUMBER: US/09/916,359  
 ; CURRENT FILING DATE: 2001-07-26  
 ; PRIOR APPLICATION NUMBER: 09/388,874  
 ; PRIOR FILING DATE: 1999-09-02  
 ; PRIOR APPLICATION NUMBER: 97/02,887  
 ; PRIOR FILING DATE: 1997-03-06  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 9379  
 ; TYPE: DNA  
 ; ORGANISM: Virus

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (320)... (9352)
US-09-916-359-1

Alignment Scores:
Pred. No.: 6,02e-166 Length: 9379
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 9 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-09-916-359-1 (1-9379)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1439 GTGCTGCTGCTATTTGCCGGCGTC-----GACGCGGAACCCACGTCACC 1483
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1484 GGGGGAAGTCCGGCCACACTGTGCTGGATTGTTAGCTCTCGCACGAGCGCCAAAG 1543
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 1544 CAGAACGTCAGCTGATCAACCAACGCGAGTTGGCACCTCAATAGCACGCGCCCTGAAC 1603
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheThrHisHisLysPheAsn 90
Db 1604 TGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCACAAGTTCAAC 1663
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1664 TCTTCAGGCTGCTCCTGAGAGGCTAGCCAGCTGCGCACCCCTTACCGATTTTGACCAAGGCG 1723
QY 111 TrpGlyProIleSerTrpAlaAsnGlySerGlyProAspGlnArgProTrpCysTrpHis 130
Db 1724 TGGGGCCCTATCAGTTATGCCAACGGAAGGGCCCGACAGGCTGTGTGGTCCGGTATATTGC 1783
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 1784 TACCCCCCAAACTTGGCGTATTGTCGCCGGAACAGAGTGTGTGGTCCGGTATATTGC 1843
QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1844 TTCACCTCCAGCCCCCTGTGTGGGAACGACGACAGTTCGGGCGCGCCACCTACAGC 1903
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1904 TGGGGTGAAATGATACGACGCTCTTCGTCCTTAAACATACACGCGCACCGCTGGGCAAT 1963
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1964 TGCTTCGGTTGTACCTGGATGAACCTCACTGGATTCAACAAAGTGTGCGGAGCGGCTCCT 2023
QY 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 2024 TGTGTCTATCGGAGGGCGGGCAACAACACCTGCACTGCGCCCACTGATTGCTTCGCGAAG 2083
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 2084 CATCCGAGCGCCACATCACTCTCGGTGGGCTCGGTCCTCGTCCATACACCCAGGTGCCTG 2143
QY 251 ValAspTyrProTrpArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 2144 GTCGACTACCGGTATAGGCTTTGGCATTTATCCTTGTACCATCAACTACACCATATTAAA 2203
QY 271 IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
Db 2204 ATCAGATGTACGTGGAGGGGTCGAACACAGGCTGGAAAGCTGCAACTGCAACGCGCG 2263
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303

```

```

Db 2264 GCGGAACGTTGCGATCTCGAAGACAGGACAGGTCGGAG 2302

RESULT 5
US-10-445-724-1
; Sequence 1, Application US/10445724
; Publication No. US20040101829A1
; GENERAL INFORMATION:
; APPLICANT: STAPLETON, JACK T.
; APPLICANT: WUENSCHMANN, SABINA
; TITLE OF INVENTION: A PROTEIN THAT INTERACTS WITH LIPIDS AND
; FILE REFERENCE: IOWA:045US
; CURRENT APPLICATION NUMBER: US/10/445,724
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: 60/392,158
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9401
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342).. (9374)
US-10-445-724-1

Alignment Scores:
Pred. No.: 6,03e-166 Length: 9401
Score: 1620.50 Matches: 283
Percent Similarity: 96.93% Conservative: 1
Best Local Similarity: 96.59% Mismatches: 4
Query Match: 92.81% Indels: 5
DB: 18 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-10-445-724-1 (1-9401)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1461 GTGCTGCTGCTATTTGCCGGCGTC-----GACGCGGAACCCACGTCACC 1505
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1506 GGGGGAAGTCCGGCCACACTGTGCTGGATTGTTAGCTCTCGCACGAGCGCAAG 1565
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 1566 CAGAACGTCAGCTGATCAACCAACGCGAGTTGGCACCTCAATAGCACGCGCCCTGAAC 1625
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheThrHisHisLysPheAsn 90
Db 1626 TGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCACAAGTTCAAC 1685
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1686 TCTTCAGGCTGCTCCTGAGAGGCTAGCCAGCTCGGACCCCTTACCGATTTCACAGGCG 1745
QY 111 TrpGlyProIleSerTrpAlaAsnGlySerGlyProAspGlnArgProTrpCysTrpHis 130
Db 1746 TGGGGCCCTATCAGTTATGCCAAGCGGCGCCCGACGAGCGGCGGCTTACTGCTGGCAC 1805
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 1806 TACCCCCCAAAACCTTGGGTATTGTGCCCGGAAGAGTGTGTGGTCCGGTATATTGC 1865
QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1866 TTCCTCCAGCCCCCTGGTGGTGGGAACGACGACAGTTCGGGCGGCGCCACCTACAGC 1925
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1925 TGGGGTGAAATGATACGAGCGCTCTTCGTCCTTAAACATACAGGCGCCACCGCTGGGCAAT 1985

```

191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
1986 TGGTTCGGTGTACCTGGATGAACCTCAACTGGATTACCAAGGTGTGGAGCGCCCTCT 2045  
211 CysValIleGlyGlyAlaGlyAsnAsnThrLysHisCysProThrAspCysPheArgLys 230  
2046 TGTGTCACTGGAGGGGGGGGCAACACACCTGCACTGCCCACTGATTCCTCCGCAAG 2105  
231 HisProAspAlaThrTrpSerArgCysGlySerGlyProThrPheThrProArgCysLeu 250  
2106 CATCCGAGCGCACATACTCTCGGTGGGCTCCGGTCCCTGGATCAACCCAGGTGCTG 2165  
251 ValAspTyrProThrArgLeuThrPheThrProCysThrLysLeuAsnThrLysPheLys 270  
2166 GTCACTACCGTATAGCTTGGCAATTATCTTGTACCACTCACTACACCATATTTAA 2225  
271 IleArgMetTyrValGlyGlyValGlyHisArgLeuGluAlaCysAsnThrThrArg 290  
2226 ATCAGGATGTACGTGGAGGGGTGCAACACAGGCTGGAAGCTGCTGCACTGGACGG 2285  
291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
2286 GCGCAACGTTGCGATCTGGAGACAGGACAGGTCCGAG 2324

RESULT 6  
US-09-742-659-3  
; Sequence 3, Application US/09742659  
; Patent No. US20010034019A1  
; GENERAL INFORMATION:  
; APPLICANT: Hong, Zhi  
; APPLICANT: Butkiewicz, Nancy J.  
; APPLICANT: Zhong, Weidong  
; APPLICANT: Ingravallo, Paul  
; APPLICANT: Wright-Minogue, Jacquelyn  
; APPLICANT: Lau, Johnson Y.  
; APPLICANT: Lemon, Stanley M.  
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses  
; FILE REFERENCE: ID01116  
; CURRENT APPLICATION NUMBER: US/09742,659  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: US 60/171,469  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 9646  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
US-09-742-659-3

Alignment Scores:  
Pred. No.: 2 56e-155 Length: 9646  
Score: 1523.50 Matches: 263  
Percent Similarity: 93.17% Conservative: 10  
Best Local Similarity: 89.76% Mismatches: 15  
Query Match: 87.26% Indels: 5  
DB: Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-09-742-659-3 (1-9646)

QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
Db 1461 GTGCTGCTGCTATTTGCGCGGTC-----GACGCGGAACCCACGTCACC 1505  
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
Db 1506 GGGGGAAGTCCCGGCCACACACCGCTGGGCTGTTGGTCTCTTACACAGGCGCAAG 1565  
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerThrPheHisLeuAsnSerThrAlaLeuAsn 70  
Db 1566 CAGAACATCCATCATGATCAACACACAGGCGAGTTGGCAGCATGACAGCGCTTGAAC 1625  
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheThrHisHisLysPheAsn 90

Db 1626 TGCAATGAAGCCCTTAACACCGGCTGGTTCAGCAGGCTCTTCTATCAGCACAAATTCAC 1685  
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
Db 1686 TCTTCAGCGTGTCTCGAGAGGTTGGCCAGCTGCGCAGCGCTTACCGATTTTGGCCAGGCG 1745  
QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130  
Db 1746 TGGGTCTCTATCAGTTATGCACCGAAGGGCTTCAGCAGNACGCCCTACTGCTGGCAGC 1805  
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
Db 1806 TACCCTCAAAGACCTTGTGCATTTGTCGCCGAAAGAGCGTGTGTGCCCGGTATATTCG 1865  
QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170  
Db 1866 TTCACCTCCAGCCCCGCTGGTGGGAACGACCGACAGGTGGGCGCGCTTACCTACAGC 1925  
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
Db 1926 TGGGTGCAATGATACGGATGCTTCGCTCTTAAACACACAGGCGCACCGCTGGGCAT 1985  
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
Db 1986 TGGTTCGGTGTACCTGGATGAACCTCAACTGGATTACCAAAAGTGTGGAGCGCCCCC 2045  
QY 211 CysValIleGlyAlaGlyAsnAsnThrLysHisCysProThrAspCysPheArgLys 230  
Db 2046 TGTGTCACTCGAGGGGTGGGCAACACACCTTGTCTGCCCCACTGATTGCTTCCGCAA 2105  
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProThrPheThrProArgCysLeu 250  
Db 2106 CATCCGGAACCCACATACTCTCGGTGGGCTCCGCTCCCTGGATTACCCAGGTGCA 2165  
QY 251 ValAspTyrProThrArgLeuThrPheThrLysThrProCysThrLysLeuAsnThrLysPheLys 270  
Db 2166 GTCGACTACCGTATAGCTTTGGCACTATCTTGTACCATCAATTAACCATATTCAAA 2225  
QY 271 IleArgMetTyrValGlyGlyValGlyHisArgLeuGluAlaCysAsnThrThrArg 290  
Db 2226 GTCAGGATGTACGTGGAGGGGTGCGAGCAGGCTGGAAGCGGCTGCAACTGGACGCG 2285  
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
Db 2286 GCGCAACGCTGCTGATCTGGAGACAGGACAGGTCCGAG 2324

RESULT 7  
US-09-238-076-19  
; Sequence 19, Application US/09238076  
; Patent No. US20020102540A1  
; GENERAL INFORMATION:  
; APPLICANT: RICE, CHARLES et al.  
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C  
; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/238,076  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US 09/034,756
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9416 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-09-238-076-19

Alignment Scores:
Pred. No.: 4,11e-155 Length: 9416
Score: 1521.50 Matches: 263
Percent Similarity: 92.83% Conservative: 9
Best Local Similarity: 89.76% Mismatches: 16
Query Match: 87.14% Indels: 5
DB: 9 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-09-238-076-19 (1-9416)

QY 11 ValLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
DB 1461 GTGCTGCTGCTATTTCGCCGGCGTC-----GACGCGGAACCCACGTCACC 1505

QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
DB 1506 GGGGGAAGTCCCGCCACACCCAGCGTGGCTTTGGTCTCTTACACGAGCGCCAAG 1565

QY 51 GlnAenValGlnLeuLeuAenThrAenGlySerTrpHisLeuAenSerThrAlaLeuAen 70
DB 1566 CAGAACTCAACTGATCAACACCAACGCGAGTTGGCACATCAATAGCAGCGCTTGAAC 1625

QY 71 CysAenAspSerLeuAenThrGlyTrpLeuAlaGlyLeuPheTyHisHisLysPheAen 90
DB 1626 TGCACGATAGCCTTACCAACCGGCTGTGTAGCAGGCTCTTCTATCGCCACAATTCAC 1685

QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAenPheAspGlnGly 110
DB 1686 TCTTCAGGCTGCTCTCAGAGGTTGGCCAGCTGCCGACGCTTACCGATTTCGCCAGGGC 1745

QY 111 TrpGlyProIleSerTyAlaAenGlySerGlyProAspGlnAenProTyCysTrpHis 130
DB 1746 TGGGGTCCCATCATGTTATGCCAACGGAAGCGGCTTGCAGCAACGCGCTACTGTGGCAC 1805

QY 131 TyrProPolysProCysGlyIleValProAlaLysSerValCysGlyProValTyCys 150
DB 1806 TACCTCCAGACCTTGTGGCATTGTCCCGCAAGAGCGTGTGCGCCGGTATATTGC 1865

QY 151 PheThrProSerProValValGlyThrThrAenArgSerGlyAlaProThrTySer 170
DB 1866 TTCACCTCCAGCCCCCGTGGTGGGAACGACGACAGGTGGCGCGCTACTACTACAGC 1925

QY 171 TrpGlyGluAenAspThrAenValPheValLeuAenAenThrArgProProLeuGlyAen 190
DB 1926 TGGGGTGCAATGATACGATGTCCTTCGTCCTTAAACAACACGAGGCCCGCTGGGCAAT 1985

QY 191 TrpPheGlyCysThrTrpMetAenSerThrGlyPheThrLysValCysGlyAlaProPro 210
DB 1986 TGGTTCGGTTGTAACCTGGATGAACCTCACTGGATTTCACCAAGTGTGGGAGCGCCCT 2045

QY 211 CysValIleGlyAlaGlyAenAenThrLeuHisCysProThrAenPheArgLys 230
DB 2046 TGTGTATCGAGGGGTGGGCAACAACACCTTCTCTGCCCACTGATTGCTTCGCAAA 2105

; APPLICATION NUMBER: US 09/034,756
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9416 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-09-995-937-19

Alignment Scores:
Pred. No.: 4,11e-155 Length: 9416
Score: 1521.50 Matches: 263
Percent Similarity: 92.83% Conservative: 9
Best Local Similarity: 89.76% Mismatches: 16
Query Match: 87.14% Indels: 5
DB: 10 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-09-995-937-19 (1-9416)

QY 11 ValLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
```

```
Db 1461 GTGCTGCTGCTATTGTCGGCGTC-----GACGGGAAACCCACGTCACC 1505
Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuAlaProGlyAlaLys 50
Db 1506 GGGGGAAGTGGCGGCACACACCGCTGGGCTTGTGCTCTTACACAGGCGCCAAG 1565
Qy 51 GlnAsnValGlnLeuLeuLeuAsnGlySerThrPheHisLeuAsnSerThrAlaLeuAsn 70
Db 1566 CAGAACATCCAACCTGATCAACACACCGCAGTTGGCACATCAATAGCAGGCGCTTGAC 1625
Qy 71 CysAsnAspSerLeuAsnThrGlyTyrPheAlaGlyLeuPheTyrHisHisLysPheAsn 90
Db 1626 TGCAACGATAGCTTACACCGGCTGGTTAGCAGGCTCTTCTATCGCCCAAAATTCAC 1685
Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1686 TCCTCAGGCTGCTCTGAGAGTTGGCAGCTGCCGAGCGCTTACCGATTGTCAGGCGG 1745
Qy 111 TrpGlyProLeuSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 1746 TGGGTCCTCCATCAGTTATGCAACCGAAGCGGCTTGACGAACGCCCTACTGTTGGCAC 1805
Qy 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 1806 TACCCTCCAAGACCTTGTGGCATTTGTCGCGCAAGAGCGTGTGTGCGCGGTATATTGC 1865
Qy 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1866 TTACTCTCCAGCCCCGGTGGTGGGAAACGACCGAGGTGGCGCGGCTTACCTACGAC 1925
Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
Db 1926 TGGGTCGCAATGATACGGATGCTTCGTCCTTAACACACCGAGCCACCGCTGGGCAAT 1985
Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1986 TGGTTCGGTGTGTACCTGGATGAATCACTAATGGATTACCAAGAGTGTGCGAGCGCCCT 2045
Qy 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 2046 TGTGTCATCGAGGGTGGGCAACACACCTTGCTGCCCCCACTGATTGCTTCCGCAAA 2105
Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTyrPheThrProArgCysLeu 250
Db 2106 CATCGGAAGCCACATACTCTCGGTGGGCTCCGCTCCGTTGGAATTACACCCAGGTGATG 2165
Qy 251 ValAspTyrProTyrArgLeuThrPheTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 2166 GTCGACTACCGGTATAGGCTTGGCACTATCTTGTACTATCAATTACACCATATTCAA 2225
Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaCysAsnThrThrArg 290
Db 2226 GTCAGGATGACGTGGAGGGTGGACACAGGCTGGAGCGGCTGCACACTGACCGCG 2285
Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 2286 GCGGAACGCTGTGATCTGGAAGACAGGACAGGTCAGGTCGAG 2324
```

## RESULT 9

```
US-09-917-563-19
; Sequence 19, Application US/09917563
; Publication No. US20030073080A1
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
```

```
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/917,563
; FILING DATE: 27-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/238,076
; FILING DATE: 26-JAN-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9416 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-917-563-19
```

## Alignment Scores:

```
Pred. No.: 4,11e-155 Length: 9416
Score: 1521.50 Matches: 263
Percent Similarity: 92.83% Conservative: 9
Best Local Similarity: 89.76% Mismatches: 16
Query Match: 87.14% Indels: 5
DB: 10 Gaps: 1
```

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-09-917-563-19 (1-9416)

```
Qy 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1461 GTGCTGCTGCTATTGTCGGCGTC-----GACGGGAAACCCACGTCACC 1505
Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1506 GGGGGAAGTGGCGGCACACACCGCTGGGCTTGTGCTCTTACACAGGCGCCAAG 1565
Qy 51 GlnAsnValGlnLeuLeuLeuAsnGlySerThrPheHisLeuAsnSerThrAlaLeuAsn 70
Db 1566 CAGAACATCCAACCTGATCAACACACCGCAGTTGGCACATCAATAGCAGGCGCTTGAC 1625
Qy 71 CysAsnAspSerLeuAsnThrGlyTyrPheAlaGlyLeuPheTyrHisHisLysPheAsn 90
Db 1626 TGCAACGATAGCTTACACCGGCTGGTTAGCAGGCTCTTCTATCGCCCAAAATTCAC 1685
Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
Db 1686 TCCTCAGGCTGCTCTGAGAGTTGGCAGCTGCCGAGCGCTTACCGATTGTCAGGCGG 1745
Qy 111 TrpGlyProLeuSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
Db 1746 TGGGTCCTCCATCAGTTATGCAACCGAAGCGGCTTGGACAGCGGCTTACTGTTGGCAC 1805
Qy 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 1806 TACCCTCCAAGACCTTGTGGCATTTGTCGCGCAAGAGCGTGTGTGCGCGGTATATTGC 1865
Qy 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170
Db 1866 TTCACCTCCAGCCCCGGTGGTGGGAAACGACCGAGGTGGCGCGGCTTACCTACGAC 1925
```

```
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProIleuGlyAsn 190
DB 1926 TGGGGTGCAAAATGATACGGATGCTTCGTCTTAACAACACACAGCGCCACCGCTGGGCAAT 1985
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
DB 1986 TGGTTCCGTTGTACCTGGATGAACCTCACTGGATTCCAAAGTGTGCGAGGCGCCCT 2045
QY 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
DB 2046 TGTGTATCATCGAGGGGTGGCAACAACACTTGTCTCTGCCCACTGATTGCTTCGCAAA 2105
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
DB 2106 CATCCGGAAGCCACATACTCTCGGTGCGCTCCGGTCCCTGGATTACACCCAGGTGCATG 2165
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
DB 2166 GTCCGACTACCGGTATAGGCTTTGGCACTATCTTGTACTATCAATTACACCATATTCAAA 2225
QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
DB 2226 GTCAGGATGTACGTGGAGGGGTGAGCACAGCTGGAAGCGGCTGCAACTGGACGCG 2285
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
DB 2286 GCGGAACGCTGTGATCTGGAAGACAGGACAGGTCGAG 2324

RESULT 10
US-10-475-989-2
; Sequence 2, Application US/10475989
; Publication No. US20040142320A1
; GENERAL INFORMATION:
; APPLICANT: CNRS
; TITLE OF INVENTION: PROCESS FOR THE REPLICATION OF THE HEPATITIS C VIRUS
; FILE REFERENCE: WO/01 AA CNR GENO
; CURRENT APPLICATION NUMBER: US/10/475,989
; CURRENT FILING DATE: 2003-10-27
; PRIOR APPLICATION NUMBER: FR 01/05732
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 9622
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-10-475-989-2

Alignment Scores:
Pred. No.: 4,22e-155 Length: 9622
Score: 1521.50 Matches: 263
Percent Similarity: 92.83% Conservativity: 9
Best Local Similarity: 89.76% Mismatches: 16
Query Match: 87.14% Indels: 5
DB: 18 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-10-475-989-2 (1-9622)
QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
DB 1461 GTGCTGCTGCTATTTCGCGCGTC-----GACGCGGAACCCACGTCACC 1505
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerIleuLeuAlaProGlyAlaLys 50
DB 1506 GGGGGAAGTCCGCGCACACACGCGTGGCTGTTGGTCTCTTACACAGCGGCAAG 1565
QY 51 GluAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
DB 1566 CAGAAATCACTGATGATCAACACACGCGGAGTTGGCATCATCAATAGCAGCGCTTGAAC 1625
QY 71 CysAsnAspSerLeuAsnThrGlyTyrPheValGlyLeuPheThrHisLysPheAsn 90
DB 9352% Conservativity: 11
```

```
DB 1626 TGCAACGATAGCTTTACCACCGCTGGTTAGCAGGCTCTTCTATGTCGCCACAAATTCAC 1685
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110
DB 1686 TCTTCAGGCTGTCTTGAGAGGTGTCAGCTGCCAGCGCTTACCGATTTCGCCAGGGC 1745
QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130
DB 1746 TGGGGTCCCATCATGTTATGCCAACGGAGCGGCTTGACGAACGCGCTTACTGTGGCAC 1805
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
DB 1806 TACCCTCCAAAGACCTTGTGGCATTTGTCCCGCAAGAGCGTGTGGCCCGGTATATTGC 1865
QY 151 PheThrProSerProValValGlyThrThrAspArgSerClyAlaProThrTyrSer 170
DB 1866 TTTCACTCCACGCGCGTGGTGGTGGAAACGACGACAGTCGGGCGCGCTACCTACAGC 1925
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190
DB 1926 TGGGGTGCAAAATGATACGGATGCTTCGTCTTAACAACACACAGCCACCGCTGGGCAAT 1985
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
DB 1986 TGGTTCCGTTGTACTGGATGAACCTCACTGGATTCAACAAAGTGTGCGAGCGCCCT 2045
QY 211 CysValIleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230
DB 2046 TGTGTATCATCGAGGGGTGGCAACAACACTTGTCTCTGCCCACTGATTGCTTCGCAAA 2105
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
DB 2106 CATCCGGAAGCCACATACTCTCGGTGCGCTCCGGTCCCTGGATTACACCCAGGTGCATG 2165
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
DB 2166 GTCCGACTACCGGTATAGGCTTTGGCACTATCTTGTACTATCAATTACACCATATTCAAA 2225
QY 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290
DB 2226 GTCAGGATGTACGTGGAGGGGTGAGCACAGCTGGAAGCGGCTGCAACTGGACGCG 2285
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
DB 2286 GCGGAACGCTGTGATCTGGAAGACAGGACAGGTCGAG 2324

RESULT 11
US-09-827-688-7
; Sequence 7, Application US/09827688
; Publication No. US20030165476A1
; GENERAL INFORMATION:
; APPLICANT: ORSON, FRANK
; APPLICANT: KINSEY, BERNA
; APPLICANT: BHOGAL, BALBIR
; TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION
; TITLE OF INVENTION: AGENTS
; FILE REFERENCE: P01949US1/10004014
; CURRENT APPLICATION NUMBER: US/09/827,688
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/195,680
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: HEPATITIS C
US-09-827-688-7

Alignment Scores:
Pred. No.: 8,68e-155 Length: 9365
Score: 1518.50 Matches: 263
Percent Similarity: 93.52% Conservativity: 11
```



```

Best Local Similarity: 89.76% Mismatches: 14
Query Match: 86.97% Indels: 5
DB: 10 Gaps: 1

US-10-715-665-7_COPY_1_303 (1-303) x US-09-827-688-7 (1-9365)

Qy 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1449 GTGCTGCTGCTATTGCGCGGTT-----GACGGGAAACCCACGTCACC 1493

Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1494 GGGGGAAGTGGCGGCCACACAGCGTGGCTGTTCCTCTCTCTTCCAGCGGCCAAG 1553

Qy 51 GlnAsnValGlnLeuLeuLeuAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 1554 CAGNACATCCACATGATCAACACACGAGGAGTGGCACATCAATAGCAGCGCTTGAAC 1613

Qy 71 CysAsnAspSerLeuLeuAsnThrGlyTrpLeuAlaGlyLeuPheTrpHisHisPheAsn 90
Db 1614 TGCATGAAGAGCCTTAACACCGGCTGGTTAGCAGGCTCTTCTATCACCACAAATTCAC 1673

Qy 91 SerSerGlyCysProGluArgGluLeuAlaSerCysArgProLeuThrAspPheAsnGly 110
Db 1674 TCTTCAGTGTGCTGAGAGGTTGGCCAGCTGGCGAGCGCTTACCGGATTTTGCCAGGGC 1733

Qy 111 TrpGlyProIleSerTrpAlaAsnGlySerGlyProAspGlnArgProTrpCysTrpHis 130
Db 1734 GGGGTCCTATCAGTACGCGCAACGAGGCGCTCGATGACGCCCTACTCTGCTGGCAC 1793

Qy 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150
Db 1794 TACCCCTCAAGACCTTGTGGCATTGTGCCCGCAAGAGCGTGTGGCCCGGTATATTGC 1853

Qy 151 PheThrProSerProValValGlyThrAspArgSerGlyAlaProThrTyrSer 170
Db 1854 TTCACCTCCAGCCCCGGTGGTGGGAACACGACAGGTGGCGCGCTTACTACAGC 1913

Qy 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnThrArgProProLeuGlyAsn 190
Db 1914 TGGGTCGCAATGATACGGATGCTTTGTCTTCTTAACACACGAGCGCTGGGCAAT 1973

Qy 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210
Db 1974 TGGTTCGGTGGCACCTGGATGAATCACTGATGATTCACCAAGTGTGGGAGCGCCCT 2033

Qy 211 CysValIleGlyAlaGlyAlaAsnThrLeuHisCysProThrAspCysPheArgLys 230
Db 2034 TGTGTATCGGAGGGGGTGGCAACACACCTTGTCTTGGCCCCACTGATTGCTTCCGCAAG 2093

Qy 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250
Db 2094 CATCCGGAAGCCACATACTCTCGGTGGCGCTCGGTCCCTGGATTACCCAGGTGATG 2153

Qy 251 ValAspTyrProTrpArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270
Db 2154 GTCGACTACCGGTATAGGCTTTGGCACTATCTTGTACCATCAATATACCAATATTCAAA 2213

Qy 271 IleArgMetTyrValGlyGlyValGluHisArgLeuGluAlaCysAsnTrpThrArg 290
Db 2214 GTCAGATGTACGTGGAGGGGTCGAGCAGAGCTGGAGCGGCTGCAACTGGAGCGGG 2273

Qy 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303
Db 2274 GCGCAAGCTGTGATCTGGNAGACAGGACAGGTCGAG 2312

```

RESULT 12

```

US-09-238-076-1
; Sequence 1, Application US/09238076
; Patent No. US20020102540A1
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C

```

```

; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/238,076
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/034,756
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-09-238-076-1

```

```

Alignment Scores:
Pred. No.: 1,92e-154 Length: 9646
Score: 1515.50 Matches: 262
Percent Similarity: 92.83% Conservative: 10
Best Local Similarity: 89.42% Mismatches: 16
Query Match: 86.80% Indels: 5
DB: 9 Gaps: 1

```

```

US-10-715-665-7_COPY_1_303 (1-303) x US-09-238-076-1 (1-9646)

Qy 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30
Db 1461 GTGCTGCTGCTATTGCGCGGCTC-----GACGGGAAACCCACGTCACC 1505

Qy 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50
Db 1506 GGGGGAAGTGGCGGCCACACGCGTGGCTGTTCCTTACACGAGCGCCAAG 1565

Qy 51 GlnAsnValGlnLeuLeuLeuAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70
Db 1566 CAGAACATCAACTGATCAACACCAACGCGAGTTGGCACATCAATAGCAGCGCTTGAAC 1625

Qy 71 CysAsnAspSerLeuLeuAsnThrGlyTrpLeuAlaGlyLeuPheTrpHisHisPheAsn 90
Db 1626 TGCATGAAGAGCCTTAACACCGGCTGGTTAGCAGGCTCTTCTATCAGCACAAATTCAC 1685

Qy 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAsnGly 110
Db 1686 TCTTCAGGCTGCTCGAGAGGTTGGCGAGCTGGCGAGCGCTTACCGATTTTGCCAGGGC 1745

Qy 111 TrpGlyProIleSerTrpAlaAsnGlySerGlyProAspGlnArgProTrpCysTrpHis 130
Db 1746 TGGGGTCTTATCATGATTATGCCAACGGAAGCGGCTCGACGAAACGCCCTTACTGCTGGCAC 1805

```



QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
Db 1806 TACCCTCCAAGACCTTGTGGCATTGTGCCGGAAGAGCGTGTGGCCCGGTATATTGC 1865  
QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170  
Db 1866 TTCACTCCAGCCCGCGTGTGTGGGAACGACGACAGTCTGGGGCGGCTACCTACAGC 1925  
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
Db 1926 TGGGGTGCAATGATACGAGTGTCTTCGTCTTAAACAACACGAGCCACCGCTGGGCAAT 1985  
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
Db 1986 TGGTTCGGTGTACCTGGATGAACCACTGGATTACCAAAAGTGTGGAGCGCCCGCT 2045  
QY 211 CysValIleGlyAlaGlyAsnAsnThrLeuLysCysProThrAspCysPheArgLys 230  
Db 2046 TGTGTCACTCGAGGGGTGGGCAACACACTTGTCTTGGCCCACTGATTGTTTCGCAAG 2105  
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
Db 2106 CATCGGAAGCCACATACCTCTCGGTGGCTCCGGTCCCTGGATTACACCCAGGTGCAG 2165  
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
Db 2166 GTCGACTACCCGCTATAGGCTTTGGCACTATCTTGTACCATCAATTACACCATATTCAA 2225  
QY 271 IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
Db 2226 GTCAAGATGTACTGGAGGGGTGCGACACAGCTGGAGCGGCTGCAACTGGAGCGG 2285  
QY 291 GlyLysArgCysAspLeuGluAspArgAspArgSerGlu 303  
Db 2286 GCGAAGCGTGTGATCTGGAAGACAGGACAGGTCCGAG 2324

RESULT 13

US-09-995-937-1  
; Sequence 1, Application US/09995937  
; Publication No. US20030028010A1  
; GENERAL INFORMATION:  
; APPLICANT: RICE, CHARLES et al.  
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/995,937  
; FILING DATE: 28-No. US20030028010A1-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/034,756  
; FILING DATE: 04-May-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 6029-4831  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 9646 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-995-937-1  
Alignment Scores:  
Pred. No.: 1.92e-154 Length: 9646  
Score: 1515.50 Matches: 262  
Percent Similarity: 92.83% Conservative: 10  
Best Local Similarity: 89.42% Mismatches: 16  
Query Match: 86.80% Indels: 5  
DB: 10 Gaps: 1  
US-10-715-665-7\_COPY\_1\_303 (1-303) x US-09-995-937-1 (1-9646)  
QY 11 ValLeuLeuLeuCysGlyAlaValPheValSerProSerAlaSerGluThrHisValThr 30  
Db 1461 GTGCTGTGCTATTTTCCGGCGCTC-----GACGCGGAACCCACGCTCACC 1505  
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
Db 1506 GGGGGAAGTGGCGGCGGACACGCGCTGGGCTTGTGTCTCTTACACGAGCGGCCAAG 1565  
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70  
Db 1566 CAGAACATCCAACTCATCAACACCAACGCGAGTTGGCACATCAATAGCAGCGCCTTGAAC 1625  
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisIleLysPheAsn 90  
Db 1626 TGCAATGAAAGCCTTAAACACCGGCTGGTTAGCAGGCTCTTCTATCAGCACAAATTC AAC 1685  
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
Db 1686 TCTTCAGGCTGTCTTGAGAGGTGGCCAGCTGCCAGCGCTTACCGATTTTGCACAGGCG 1745  
QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130  
Db 1746 TGGGGTTCCTATCAGTTATGCCAACGGAAGCGGCTCGACGAACGCGCTTACTGTGGCAC 1805  
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
Db 1806 TACCCTCCAAGACCTTGTGGCATTGTGCCGGAAGAGCGTGTGGCCCGGTATATTGC 1865  
QY 151 PheThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTyrSer 170  
Db 1866 TTCCTCCAGCCCGCGTGTGTGGGAACGACGACAGTCTGGGCGGCGCTACCTACAGC 1925  
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
Db 1926 TGGGGTGCAATGATACGAGTGTCTTCGTCTTAAACAACACGAGCCACCGCTGGGCAAT 1985  
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
Db 1986 TGGTTCGGTGTACCTGGATGAACCACTGGATTACCAAAAGTGTGGAGCGCCCGCT 2045  
QY 211 CysValIleGlyAlaGlyAsnAsnThrLeuLysCysProThrAspCysPheArgLys 230  
Db 2046 TGTGTCACTCGAGGGGTGGGCAACACACTTGTCTTGGCCCACTGATTGTTTCGCAAG 2105  
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
Db 2106 CATCGGAAGCCACATACCTCTCGGTGGCTCCGGTCCCTGGATTACACCCAGGTGCATG 2165  
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
Db 2166 GTCGACTACCCGCTATAGGCTTTGGGCACTATCTTGTACCATCAATTACACCATATTCAA 2225  
QY 271 IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290

Db 2226 GTCCAGGATGACGTGGAGGGGTCCAGCACAGGCTGCAACTGGACGGG 2285  
QY 291 GlyGluArgCysAspLeuGluAlaPheValSerProSerAlaValThrHisValThr 303  
Db 2286 GCGCAACGCTGTGATCTGGAGAGACAGGACAGGTCCGAG 2324

## RESULT 14

US-09-917-563-1  
; Sequence 1, Application US/09917563  
; Publication No. US20030073080A1  
; GENERAL INFORMATION:

; APPLICANT: RICE, CHARLES et al.  
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C  
; VIRUS (HCV) AND USES THEREOF

; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/917,563  
; FILING DATE: 27-Jul-2001

; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/238,076  
; FILING DATE: 26-JAN-1999

; ATTORNEY/AGENT INFORMATION:  
; NAME: HOLLAND, DONALD R.  
; REGISTRATION NUMBER: 35,197  
; REFERENCE/DOCKET NUMBER: 6029-4831  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 314-727-5188  
; TELEFAX: 314-727-6092  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9646 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-917-563-1

## Alignment Scores:

Pred. No.: 1.92e-154 Length: 9646  
Score: 1515.50 Matches: 262  
Percent Similarity: 92.83% Conservative: 10  
Best Local Similarity: 89.42% Mismatches: 16  
Query Match: 86.80% Indels: 5  
DB: 10 Gaps: 1

US-10-715-665-7\_COPY\_1\_303 (1-303) x US-09-917-563-1 (1-9646)

QY 11 valLeuLeuLeuCysGlyAlaValPheValSerProSerAlaValThrHisValThr 30  
Db 1461 GTCTGCTGCTATTTCGCCGCTC-----GACCGGAAACCCACGTCACC 1505  
QY 31 GlyGlySerAlaGlyHisThrValSerGlyPheValSerLeuLeuAlaProGlyAlaLys 50  
Db 1506 GGGGGAGTCCCGCGCCACACAGGCTGGGCTGTTCCTTACACAGCGCGCAAG 1565  
QY 51 GlnAsnValGlnLeuIleAsnThrAsnGlySerTrpHisLeuAsnSerThrAlaLeuAsn 70

Db 1566 CAGAACATCAACTGATCAACACCAACGCGAGTTGGCACATCAATAGACGCGCTTGAAC 1625  
QY 71 CysAsnAspSerLeuAsnThrGlyTrpLeuAlaGlyLeuPheTyrHisHisLysPheAsn 90  
Db 1626 TGCATGAAAGCCCTTAACACCGGCTGGTTAGCAGGGCTCTTCTATCAGCACAAATTCAC 1685  
QY 91 SerSerGlyCysProGluArgLeuAlaSerCysArgProLeuThrAspPheAspGlnGly 110  
Db 1686 TCCTCAGGCTGTCTTGAGAGGTTGGCCAGCTGCCAGCGCTTACCGATTTTCCAGGGC 1745  
QY 111 TrpGlyProIleSerTyrAlaAsnGlySerGlyProAspGlnArgProTyrCysTrpHis 130  
Db 1746 TGGGGTCTCTATCAGTTATGCAACGGAAGCGGCTCGACGAAACGCCCTCTACTGCTGCAC 1805  
QY 131 TyrProProLysProCysGlyIleValProAlaLysSerValCysGlyProValTyrCys 150  
Db 1806 TACCTCCAAAGCCTTGTGGCATTTGTCCCGCAAGAGCGTGTGTGCCGGTATATATGC 1865  
QY 151 PheThrProSerProValValGlyThrAspArgSerGlyAlaProThrTyrSer 170  
Db 1866 TTCACTCCACAGCCCCGTGGTGGTGGNACACGACAGGTGGGGCGGCTACCTACAGC 1925  
QY 171 TrpGlyGluAsnAspThrAspValPheValLeuAsnAsnThrArgProProLeuGlyAsn 190  
Db 1926 TGGGGTGCATAATGATACGGATGTCTTCGTCTTTAAACAAACACACGAGCCACCGCTGGCAAT 1985  
QY 191 TrpPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProPro 210  
Db 1986 TGGTTCCGTTGTACCTGGATGAACTCAACTGGATTTACCAAGTGTCCGAGCGCCCCCT 2045  
QY 211 CysValIleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
Db 2046 TGTGTCTATCGAGGGGTGGGCAACACACCTTGTCTGCCCCACTGATTGTTTCCGCAAG 2105  
QY 231 HisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrProArgCysLeu 250  
Db 2106 CATCCGAGACCCACATACTCTCGGTGGTCCGCTCGGTCCCTGGATTACACCGAGTGCATG 2165  
QY 251 ValAspTyrProTyrArgLeuTrpHisTyrProCysThrIleAsnTyrThrIlePheLys 270  
Db 2166 GTCGACTACCCGTATAGGCTTTGGCACTATCCTTGTACCATCAATATACCATATTCAA 2225  
QY 271 IleArgMetTyrValGlyValGluHisArgLeuGluAlaAlaCysAsnTrpThrArg 290  
Db 2226 GTCAGGATGTACGTGGAGGGGTCCGACACAGGCTGGAAGCGGCTGCACTGGACCGG 2285  
QY 291 GlyGluArgCysAspLeuGluAspArgAspArgSerGlu 303  
Db 2286 GCGCAACGCTGTGATCTGGAAGACAGGACAGGTCGAG 2324

## RESULT 15

US-09-238-076-5

; Sequence 5, Application US/09238076  
; Patent No. US20020102540A1  
; GENERAL INFORMATION:

; APPLICANT: RICE, CHARLES et al.  
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C  
; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
; STREET: 7733 FORSYTH BLVD., SUITE 1400  
; CITY: ST. LOUIS  
; STATE: MO  
; COUNTRY: USA  
; ZIP: 63105  
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 15, 2005, 04:37:38 ; Search time 3682 Seconds  
(without alignments)  
3132.394 Million cell updates/sec

Title: US-10-715-665-7\_COPY\_1\_303

Perfect score: 1746

Sequence: 1 MDAMKRGCLCCVLLCCGAVFV.....AACNWTGRGCDLEDNRDRSE 303

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US10715665/runat\_13042005\_165823\_28155/app\_query.fasta\_1.455  
-DBS=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0  
-UNITS=bite -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-NSR=US10715665 @CGN 1 1 3437 @runat\_13042005\_165823\_28155 -NCPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	171	9.8	899	2	BE277041
2	164	9.4	904	5	BUS0108
3	159.5	9.1	716	2	BE408331
4	159.5	9.1	835	4	BG281371
5	157.5	9.0	761	4	BI771956
6	157	9.0	899	5	BQ890669
7	157	9.0	913	4	BG476931
8	154.5	8.8	798	4	BG281523
9	154.5	8.8	1041	5	BQ072488

10	154	8.8	511	2	BE389162
11	153	8.8	1050	2	BE312788
12	153	8.8	1124	4	BM553317
13	150.5	8.6	860	5	BQ682920
14	150.5	8.6	893	6	CB992325
15	150	8.6	951	5	BQ687808
16	148.5	8.5	898	4	BI757143
17	148.5	8.5	969	1	AL550289
18	147.5	8.4	897	2	BF311265
19	147	8.4	904	5	BU157553
20	146.5	8.4	877	1	AU131845
21	146	8.4	967	5	BX449214
22	146	8.4	1034	2	BE314639
23	145.5	8.3	694	1	AL524377
24	145.5	8.3	942	5	BQ880098
25	144.5	8.3	797	7	CF994970
26	143.5	8.2	877	5	BUS57319
27	143.5	8.2	883	6	CD359959
28	143.5	8.2	1029	4	BG478933
29	143	8.2	928	5	BU149873
30	143	8.2	1174	4	BG253421
31	142.5	8.2	805	7	CF132126
32	142	8.1	817	1	AU116939
33	142	8.1	821	5	BU173689
34	142	8.1	888	5	BQ689818
35	141.5	8.1	874	4	BG334653
36	141.5	8.1	907	1	AU131208
37	141.5	8.1	952	5	BUS38355
38	141	8.1	943	4	BM804838
39	140.5	8.0	646	4	BG477928
40	140.5	8.0	711	4	BG476511
41	140.5	8.0	793	1	AU131736
42	140.5	8.0	804	2	BE736974
43	140.5	8.0	2498	3	CR623908
44	140	8.0	828	4	BG762262
45	140	8.0	845	5	BQ877849

ALIGNMENTS

RESULT 1  
BE277041  
LOCUS 601178928F1 NIH\_MGC\_20 Homo sapiens cDNA clone IMAGE:3543990 5',  
DEFINITION mRNA sequence.  
ACCESSION BE277041  
VERSION BE277041.1 GI:9152005  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 899)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LLCW238 row: h column: 07  
High quality sequence stop: 688.

source  
1. .899  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3543990"  
/tissue\_type="melanotic melanoma"

FEATURES  
Location/Qualifiers  
1. .899  
High quality sequence stop: 688.





```

Db      516 GGAACAGCAGCGCT-----TGGCCCAAGAAGCCCACA   548
Qy      157 ValValGlyThrThrAspAgsSerGlyAlaProHrrTySertTpGLYGIUasNAspThr    176
          |||              |||||             :|||||           |||
Db      549 CGCGGCAGGCCAGCACCATCAGGC-----TGGCCTGGGAAACCACACTACT     599
Qy      177 AspVal-----PheValLeuAsnAnstThrArpPro        186
          ::               |:|||            |||              |||
Db      600 GAGAACCCAGATCGAGACTCAAGCCTGTGTCTACGTTTATAAGCGGGAAATACACGTCCTCA  659
Qy      187 -----ProLeuGliYAasnPhpeGlyCysThrTrpMet       197
          :|||         :|||           |||
Db      660 AGTTGCCAACCCCTCTCTCTTGCGGACGGCTGTCTGGATG   701


RESULT 4
BG281371                                835 bp mRNA linear EST 21-FEB-2001
LOCUS 602401866F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4544200 5' ,
DEFINITION mRNA sequence.
ACCESSION BG281371 GI:13030296
VERSION   BG281371.1
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
REFERENCE 1 (bases 1 to 835)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: ATCC/DCTD/DTIC CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov Plate: LHCMI227 row: C column: 17 High quality sequence stop: 826. Location/Qualifiers
FEATURES
source
1 . 835
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
clone="IMAGE:4544200"
tissue_type="melanotic melanoma"
lab_hosts="DHIOB (phage-resistant)"
clone_lib="NIH_MGC_20"
note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
Pred.No.:                Length:                835
Score:                   Matches:                 68
                        Mismatch:                  16
Best Local Similarity:   Conserved:                 86
Query Match:             Indels:                    59
DB:                      Gaps:                       4

US-10-715-665-7_COPY_1_303 (1-303) x BG281371 (1-835)

Qy      1 MetaSpalaMetLysArggLYeucCysValLeuleuCysGYlAlaVAlpheVal    20
          |||
Db      84 ATGGATGCAATTGAAGAGAGGGCTCTGCTGTGTCTGTGTGTGGACAGTCTTCGTT   143
```



```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5208117"
/lab_host="DH10B"
/clone_lib="NIH_MGC_122"
/notes="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH MGC Library."

```

## ORIGIN

Alignment Scores:		
Pred. No.:	0.00132	Length:
Score:	157.50	Matches:
Percent Similarity:	33.47%	Conservative:
Best Local Similarity:	26.03%	Mismatches:
Query Match:	9.02%	Indels:
DB:	4	Gaps:
		11
US-10-715-665-7 COPY 1 303 (1-303) x BI771956 (1-761)		761

US-10-715-665-7 COPY 1 303 (1-303) X BI771956 (1-761)

	Qy	1	MetAspAlaMetLysArgGlyLeuCysCysValleuleuLeuCysGlyAlaValPheVal	20
	Dd	88	ATGGATGCATAAGAGAGGGCTCGTGTGTGCTGTGCTGTGTGTGAGCAGTCCTTCGTT	147
	Qy	21	SerProSerAlaSerGluThrHisVal-----	29
	Dd	148	TGCCCCAGC---CAGAANAATCCNGCCCGATTACAGAGAGACCAGATCTTACCAAGT	204
	Qy	30	-----ThrGlyGlySerAlaGlyHisThrVal-----SerGly	40
	Dd	205	TGCACGACGACCAAGTGTTTTCAAAGGGGGCACCTGCCACAGAGCCCTGTACTTCTCAGAT	264
	Qy	41	PheValSerLeuAlaProGlyAlalysGlnAenVal-----	53
	Dd	265	TTCTGTGTCCAGTGCCTCCCAGAGGATTTGCTGGGAAGTGCTGGAATAGATACACAGGGCC	324
	Qy	54	-----GlnLeuIlaAenThrAasnGlySertTrpHisLeuAenSerThrAla	68
	Dd	325	ACGTGCTACGAGGACCATGC-ATCAGCTACAGGGCCACGTGGAGCACAGCGGAGAGTGCG	383
	Qy	69	LeuAanCyAAenAaspSerLeuAanThrClYTrpLeuAlaGlyLeuPheTyRHisHISlys	88
	Dd	384	GCGGAGTGC-----ACNACTGG-----	401
	Qy	89	PheAnSerSerGlyCyAProGluAlaLeuAlaserCyArqProLeuthrAspPheAasp	108
	Dd	402	--AACAGCAGCGCGTTGGCCAGAGAAGCCCTACAGCGGGCGGAGGCAGATGCCATCAGG	458
	Qy	109	GInGlyTrpGlyProIleserTyRAalnGlySerGlyProasglnArgProTyCys	128
	Dd	459	CTGGGCCCTGGGGGAACCAACAATCTGCAGAAAACCAGATCGAGACTCAAAGCCCTGTGC	518
	Qy	129	TrpHisTyRProProLyseProCySGlyileValProAlalyserValCysGlyProval	148
	Dd	519	TAGTCTTT-----AAGCGGGGAAGTACAGCTCAGAG-----	551
	Qy	149	TyrCysPheThrProSerProVaivalGlylYthrThrAspArgSerglyAlaProthr	168
	Dd	552	TTCTGCAGCACCCCT-	566
	Qy	169	TyrSerTrpGlyGluaAenAaspThrAaspValpheValleuAenAenthArgProproLeu	188
	Dd	567	---GCCTGCTCTGAGGGNACAGTGTGCTAC-----	599
	Qy	189	GlyAanTrpPheGlyCyethrTrpMetAasnSerThrGlyPheThrLysValCysGlyAla	208

Db	600	GGGAACGGGTGAGCCTACGCTGGCAGCACAGACT---				linear	EST 16-AUG-2002
Qy	209	ProPro 210					
Db	657	CCTCCC 662					
RESULT 6							
LOCUS	BQ890669						
DEFINITION	AGENCYCOURT 8152227 Lupski dorsaal root_ganglion Homo sapiens CDNA						
	clone IMAGE:6179899 5', mRNA sequence.						
ACCESSION	BQ890669						
VERSION	BQ890669.1	GI:22282683					
KEYWORDS	EST.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE	1 (bases 1 to 899)						
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .						
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)						
JOURNAL	Unpublished (1999)						
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgabbs@mail.nih.gov">cgabbs@mail.nih.gov</a> Tissue Procurement: Dr. James R. Lupski CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM13561 row: m column: 20 High quality sequence stop: 587.						

## FEATURES

```

location/Qualifiers
1..899
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6179899"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupeksi dorsal root ganglion"
/vector="pCMT-SORT6 (Life Technologies)"
Site_1:
NotI; Site2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCACGCGTCCG-3' and
5'-GACTAGATTTCAGTCAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupeksi, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

```

## ORIGIN

Alignment Scores:

Pred. Score:	0.00185	Length:	899
Score:	157.00	Matches:	75
Percent Similarity:	28.62%	Conservative:	14
Best Local Similarity:	24.12%	Mismatches:	84
Query Match:	8.99%	Indels:	138
DB:	5	Gaps:	15

US-10-715-665-7 COPY 1 303 (1-303) x B08906569 (1-899)

[illegible]





Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1041)  
 NIH-MGC http://mgs.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM12810 row: k column: 23  
 High quality sequence stop: 625.  
 Location/Qualifiers  
 1..1041  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5761678"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 122"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;  
 Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source  
 anonymous pool of 24 week female lung, 16 week female  
 spleen, and 20-22 week male spleens. Library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.4 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 026. Note:  
 this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:  
 Pred. No.: 1041  
 Score: 154.50  
 Percent Similarity: 34.89%  
 Best Local Similarity: 28.42%  
 Query Match: 8.85%  
 DB: 5  
 Gaps: 18

US-10-715-665-7\_COPY\_1\_303 (1-303) x BQ072488 (1-1041)

Qy 1 MetAspAlaMetLysArgGlyLeuCysValLeuLeuLeuCysGlyAlaValPheVal 20  
 Db 196 ATGGATGCANTGAGAGAGGGCTCTGCTGTGCTGTGCTGTGAGCAGCTTCGTT 255  
 Qy 21 SerProSerAlaSerGluThrHisVal-----ThrGlyGlySerAlaGlyHisThrVal 38  
 Db 256 TCGCCAGC---CAGGAATTCATGCCCCGATTTCAGAAAGAGGAGCCAGATCTTACCAAGTG 312  
 Qy 39 -----SerGlyPheValSerLeuLeuAlaPro 47  
 Db 313 ATCTGCAGATGAAACCGCAGATGATATACGACCAACATCAGTCAGTGGCTGCCCT 372  
 Qy 48 GlyAlaLysGlnAsnValGlnLeulleAsnThrAsnGlySerTrpHisLeuAsnSerThr 67  
 Db 373 GTGCTCAGAAGCAC-----CGGTGGAATATTGCTGG---TGCAACAGTGGC 417  
 Qy 68 AlaLeuAsnCysAsnAsp-----SerLeuAsnThr 77  
 Db 418 AGGGCAGTGGCCACTCAGTGCCCTGTCAAAGTTGCAGCGAGCCCAAGGTGTTTCAACGGG 477  
 Qy 78 GlyTrpLeuAlaGlyLeuPheThrHisHisLysPheAsnSerSerGlyCysProGluArg 97  
 Db 478 GGCACTTCGCAGAGGCCCTGACTTCTCAGATTTCGTGTGCCAG---TGCCCCGAGGA 534  
 Qy 98 LeuAla-----SerCysArgProLeuThrAspPheAspGlnGlyTrp 111  
 Db 535 TTGTCTGGGAGTGTGTGAATAGATACACGGGCCCGCTGCTACGAGGACCGAGGC--- 591

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

112 GlyProIleSerTyrAlaAsnGlySer-GlyProAspGlnArgProTyrCysTrpHisty 131  
 Db 592 -----ATCAGCTAC---AGGGGACGCTGGAGCAGCGGAGA----- 625  
 Qy 131 rProProLysProCysGlylleValProAlaLysSerValCysGlyProValTyrCysPhe 151  
 Db 626 -----GTGGCGCCGAGTGCCACCAACTGGAACAGCAGCGCGTTG 663  
 Qy 151 eThrProSerProValValValGlyThrThrAspArgSerGlyAlaProThrTrpSerTr 171  
 Db 664 GCCCAGAAGCCCTACAGCGCGGAGGCGCAGATGCCATCAGGC-----TGGCGCTG 714  
 Qy 171 pGlyGluAsnAspThrAspValPheValLeuAsnThrArgProProLeuGlyAsnTr 191  
 Db 715 GGGAAACAACAATCT-----GCAGAAACAGATCGAGACTCAGACCCCTG 759  
 Qy 191 pPheGlyCysThrTrpMetAsnSerThrGlyPhe---ThrLysValCysGlyAlaProPr 210  
 Db 760 G-----TGCTACGCTCTTTTAGCGGGGAGTACAGCTCAGAGTCTGCAGCACCCTGC 813  
 Qy 210 cCysValleGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLys 230  
 Db 814 CTGC-----TCTGAGGGAAAC-----AGTGAATGCTACTC----- 841  
 Qy 230 sHisProAspAlaThrTyrSerArgCysGlySerGlyProTrpIleThrPro 247  
 Db 842 -----TTTGGGGAATGGGTACGCCCT 862

RESULT 10  
 BE389162 511 bp mRNA linear EST 21-JUL-2000  
 LOCUS 601285961F1 NIH\_MGC\_44 Homo sapiens cDNA clone IMAGE:3607767 5',  
 DEFINITION mRNA sequence.  
 BE389162  
 BE389162.1 GI:9334527  
 ACCESSION  
 VERSION  
 KEYWORDS  
 EST.  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 511)  
 AUTHORS  
 NIH-MGC http://mgs.nci.nih.gov/.  
 TITLE  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL  
 Unpublished (1999)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCW259 row: i column: 16  
 High quality sequence stop: 510.  
 Location/Qualifiers  
 1..511  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3607767"  
 /tissue\_type="endometrium, adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_44"  
 /note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5',  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

FEATURES  
 source

ORIGIN



QY 176 rAspValPheValLeuAsnAsnThrArgProProLeu-----GlyAsnTr 191  
 Db 667 CGAGGTTCTG-----CAGACCTGCTCTCTGAGGGGAACCGTACTGGTAAC-- 715  
 QY 191 pPheGlyCysThrTrpMetAsnSerThrGlyPheThrLysValCysGlyAlaProProCy 211  
 Db 716 -TTTGGG--GATTGGGTGAGTACCGGGG--ACGCACAGCTCCCCGAGTGGGTG 768  
 QY 211 sValleGlyGlyAlaGlyAsnAsnThrLeuHisCysProThrAspCysPheArgLysHi 231  
 Db 769 CCTCTGACTTCCCGGGGAATTCCCTGATACCTGTTAGCGCAGCGGTTTC----- 820  
 QY 231 sProAspAlaThrTySer-----Ar 238  
 Db 821 -CCGGCGCG--TATAGCCCCCGTCTCCGGCAGCTGGCTTGGCGCACCTTATT 876  
 QY 238 cGysGlySerGlyPro-----TrrPheThrProArg 248  
 Db 877 GATGGGGGGGACCGTACTGGGGTGGCTCACACCCCGT 916

RESULT 12  
 BM553317  
 LOCUS  
 DEFINITION AGENCOURT\_6542530 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5742949  
 5', mRNA sequence.  
 ACCESSION BM553317  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM12761 row: 0 column: 14  
 High quality sequence stop: 756.  
 Location/Qualifiers  
 1. .1124  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5742949"  
 /tissue\_type="medulla"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_119"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: EcoRV (destroyed); RNA source normal medulla from  
 anonymous male age 27. Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.3 kb, insert size range  
 0.9-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 013. Note:  
 this is a NIH\_MGC Library."

FEATURES  
 source  
 1. .1124  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5742949"  
 /tissue\_type="medulla"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_119"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: EcoRV (destroyed); RNA source normal medulla from  
 anonymous male age 27. Library is oligo-dT primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.3 kb, insert size range  
 0.9-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 013. Note:  
 this is a NIH\_MGC Library."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.00581 Length: 1124  
 Score: 153.00 Matches: 81  
 Percent Similarity: 30.23% Conservative: 26  
 Best Local Similarity: 22.88% Mismatches: 108  
 Query Match: 8.76% Indels: 139

DB: 4 Gaps: 14  
 US-10-715-665-7\_COPY\_1\_303 (1-303) x BM553317 (1-1124)  
 QY 1 MetAspAlaMetLysArgGlyLeuCysCysValLeuLeuLeuCysGlyAlaValPheVal 20  
 Db 87 ATGGATCAATGAAGAGAGGGCTCTGCTGTGTGTCTGTCTGTGTGGAGCAGCTCTCGT 146  
 QY 21 SerProSerAlaSerGluThrHisVal----- 29  
 Db 147 TCGCCAGC--CAGGAAATCCATGCCCGGATTTCAGAGAGGAGCCAGATCTTACCAAGTG 203  
 QY 29 ----- 29  
 Db 204 ATCTGCAGAGATGAAAAACCGCAGATGATATACCGAACATCATGCTATGGTGGCCCT 263  
 QY 30 -----ThrGlyGlySerAlaGlyHis 36  
 Db 264 GTGCTCAGAGCAACCGGGTGAATATTGCTGGTCAACAGTGCAGGCGCAGTGCCAC 323  
 QY 37 ThrVal----- 38  
 Db 324 TCAGTGCCTGTCAAAAGTTGCAGCGAGCAAGGTGTTTCAACGGGGGCACCTGCCAGCAG 383  
 QY 39 -----SerGlyPheValSerLeuLeuAlaProGlyAlaLysGlnAsnVal--- 53  
 Db 384 GCCCTGTACTTCTCAGATTGCTGCCAGTGCCTCCCGAAGGATTTGCTGGGAGTGTCT 443  
 QY 54 -----GlnLeuLeuAsnThrAsnGlySerTrp 62  
 Db 444 GAAATAGATACCGAGGCCACGTGTACGAGGACACGAGGCATCAGCTACAGGGGCACGTGG 503  
 QY 63 HisLeuAsnSerThrAlaLeuAsnCysAsnAspSerLeuAsnThrGlyTrpLeuAlaGly 82  
 Db 504 AGCAGCGGAGAGTGGCGCCGAGTGC-----ACCACTGG----- 539  
 QY 83 LeuPheTyrHisHisLysPheAsnSerSerGlyCysProGluArgLeuAlaSerCysArg 102  
 Db 540 -----AACAGCAGCGGTGGCCAGAACCCCTACAGCGGCGG 578  
 QY 103 ProLeuThrAspPheAspGlnGlyTrpGlyProIleSerTyrrAlaAsnGlySerGlyPro 122  
 Db 579 AGGCAGATGCCATCAGCTGGCGCTGGGAAACCACTACTCTGAGAAACCCAGATCGA 638  
 QY 123 AspGlnArgProTyrrCysTrp-----HisTyrrProProLysProCysGly 137  
 Db 639 GACTCAAGCCCTGTGTCTACGTCTTAAAGCGGGGAAGTACAGCTCAGAGTTCTGCAGC 698  
 QY 138 Ile-ValProAlaLysSerValCysGlyProValTyrCysPheThrProSerProValVa 157  
 Db 699 ACCCTGCTCTCTGAGGGAACAGTACTCTACTTTGGGAATGGGTACACCTACCT 758  
 QY 157 lValGlyThrThrAspArgSerGlyAlaProThrTyrrSerTrpGlyGluAsnAspThrAs 177  
 Db 759 GGCAGCAGACGCTCAGCGTGGGTGCTCTGCTCCCGTGAATTCATGATCCTCA 818  
 QY 177 pValPheValLeuAsnAsnThrArg-----ProProLeuGlyAsnTrpPhe----- 192  
 Db 819 TAGGCAAGGGTTTACAGCAGACAGAACCCAGTGCCTCCCGTGGCCCTGGGCAACAT 878  
 QY 193 -----GlyCys--ThrTrp-----MetAsnSerTh 200  
 Db 879 AATTACTCCCGGAATCTGTATGGGGATGCCAGCCCTGGTCCAGTGTCTGAGAACCC 938  
 QY 200 rGlyPheThr-----LysValCysGlyAlaProProCysValIleGlyGlyAlaGlyAs 218  
 Db 939 AGGNGCTGACCTGGGAANTACTGTGATGTCCTCTCTGCTCCCTTGGGNC----- 990  
 QY 218 nAsnThrLeuHisCysProThrAspCysPheArgLysHisProAspAlaThrTyrrSer-- 237  
 Db 991 -----TTGAGCAGTACAGCCAGGCGCTCATTTTCCCA 1022  
 QY 238 -----ArgCysGlySerGlyPro 243

**FEATURES**  
**source**





```
Db 199 ATCTGCAGAGATGAAAAAACCCAGATGATATACCAGCAACATCAGTCATGGCTCGGCCCT 258
Qy 30 -----ThrGlyGlySerAlaGlyHis 36
Db 259 GTGCTCAGAGCAACCGGTGGATATTTGTGTGTGCAACAGTGGCAGGGGCACAGTGCAC 318
Qy 37 ThrVal-----38
Db 319 TCAGTGCCTGTCAAAGTTGCAGCGAGCCAGAGTGTTCACGCGGGGCACCTGCCAGCAG 378
Qy 39 -----SerGlyPheValSerLeuAlaProGlyAlaLysGlnAsnVal---53
Db 379 GCCCTGACTTCTCAGATTTCGTGTGCCAGTGTGCCCGAAGGATTTGTGGGAAGTGTGT 438
Qy 54 -----GlnLeuIleAsnThrAsnGlySerTyr 62
Db 439 GAATATAGATACCAGGGCCACGTGTACAGGACCAGGGGCATCAGCTACAGGGGCACGTGG 498
Qy 63 HisLeuAsnSerThrAlaLeuAsnCysAsnAspSerLeuAsnThrGlyTyrLeuAlaGly 82
Db 499 AGCACAGCGGAGTGGCGCGAGTGC-----ACCAACTGG-----534
Qy 83 LeuPheTyrHisHisLysPheAsnSerSerGlyCysProGluArgLeuAlaSerCysArg 102
Db 535 -----AACAGCAGCGCGTGTGCCCGCAGAGCCCTACAGCGGCGG 573
Qy 103 ProLeuThrAspPheAspGlnGlyTyrGlyProIleSerTyrAlaAsnGlySerGlyPro 122
Db 574 AGCCAGACGCCATCAGCTGGGCTGGGNAACCACTACTGCAGAAACCCAGATCGA 633
Qy 123 AspGlnArgProTyrCysTyr-----HisTyrProProLysProCysGly 137
Db 634 GACTCAAAGCCCTGTGTCTACGTCTTTAAGCGGGGAAGTACAGCTCAGAGTTCTGCAGC 693
Qy 138 IleValProAla-----LysSerValCys-----GlyProValTyr 149
Db 694 ---ACCCCTGCCTGCTCTGAGGGNAACAGTGACTGCTACTNNGNAAATGGGTGAGCCTAC 750
Qy 150 CysPheThr-----ProSerProValValGlyThr 160
Db 751 CGTGGCAGGCACAGCCTCAACCGAGTCGGGTGCTCCTCGCTCCGTGGGAAATTCATGAT 810
Qy 161 ThrAspArgSerGly-----AlaPro-----ThrTyrSerTyr 171
Db 811 CCTGATAGGCAAGGTTTACACAGCACAAAGAACCCCAANTGCCCGCAGGCACTGNGGCCTGG 870
Qy 172 GlyGluAsn 174
Db 871 GGCAAAACAT 879
```

Search completed: April 15, 2005, 09:06:03  
Job time : 3692 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 17:35:47 ; Search time 6936 Seconds  
(without alignments)  
11128.771 Million cell updates/sec

Title: US-10-715-665-6\_COPY\_1992\_3584  
Perfect score: 1593  
Sequence: 1 atggatgcaatgaagagagg.....ttgtctctgggtatacatt 1593

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	1593	100.0 5882 6	AX154501 Sequence
2	836.4	52.5 1989 6	BD132859 Intracell
3	836.4	52.5 1989 6	AR282014 Sequence
4	834.8	52.4 1207 6	AR004336 Sequence
5	834.8	52.4 1207 6	AR006827 Sequence
6	834.8	52.4 8316 6	AR118703 Sequence
7	834.8	52.4 9185 6	AR118722 Sequence
8	834.8	52.4 9185 6	AR118723 Sequence
9	834.8	52.4 9185 6	108294 Sequence 1
10	834.8	52.4 9185 6	BD091382 HCV culti
11	834.8	52.4 9379 6	AR166930 Sequence
12	834.8	52.4 9379 6	AR301300 Sequence
13	834.8	52.4 9401 6	AR176483 Sequence
14	834.8	52.4 9401 6	E66593 Hepatitis C
15	834.8	52.4 9401 6	171894 Sequence 9
16	834.8	52.4 9401 6	181885 Sequence 9
17	834.8	52.4 9401 6	BD080334 Hepatitis
18	834.8	52.4 9401 14	M62321 Hepatitis C
19	833.2	52.3 1989 6	AR179260 Sequence

20	833.2	52.3	8987	6	AR118728
21	833.2	52.3	9609	12	AF387805
22	833.2	52.3	9609	12	AF387808
23	833.2	52.3	9646	12	AF387806
24	833.2	52.3	9693	12	AF387807
25	832.8	52.3	9379	6	AR118747
26	831.6	52.2	1210	6	AR031233
27	831.6	52.2	1210	6	AR145049
28	818.8	51.4	9618	14	AF271632
29	748.4	47.0	2610	14	HPCST77
30	745.2	46.8	9646	6	AR110828
31	745.2	46.8	9646	6	BD069982
32	745.2	46.8	9646	14	AF009606
33	745.2	46.8	12980	6	AR110831
34	745.2	46.8	12980	6	BD069985
35	743.6	46.7	9599	14	AF011753
36	742	46.6	9599	6	AR119831
37	742	46.6	9599	14	AF011751
38	742	46.6	12119	6	AX057315
39	740.4	46.5	5125	6	AR068510
40	740.4	46.5	5125	6	165457
41	740.4	46.5	5323	6	AR068509
42	740.4	46.5	5323	6	165456
43	740.4	46.5	7551	6	CQ795336
44	740.4	46.5	7551	6	CQ867240
45	740.4	46.5	7754	6	CQ785590

ALIGNMENTS

RESULT 1  
AX154501  
LOCUS AX154501 5882 bp DNA linear PAT 22-JUN-2001  
DEFINITION Sequence 6 from Patent WO0138358.  
ACCESSION AX154501  
VERSION AX154501.1 GI:14536111  
KEYWORDS synthetic construct  
SOURCE other sequences; artificial sequences.  
ORGANISM 1  
REFERENCE Selby, M.C., Glazer, E.C. and Houghton, M.C.  
AUTHORS Hbv/hcv virus-like particle  
TITLE Patent: WO 0138358-A 6 31-MAY-2001;  
JOURNAL CHIRON CORPORATION (US)  
FEATURES Location/Qualifiers  
source  
1..5882  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="plasmid pCMV-II-E2661-sag"  
1992..3587  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAC42602.1"  
/db\_xref="GI:14536112"  
/translation="MDAMKRGCCVLLCGAVFSPASETHVTGSGAGTSGFVSL  
LAPGAKQVQLINTGSHWLNALNCNSLNTGLAGLFYHHKFNSSGCPRLASCR  
PLTDFQGWGPISYANGSDQRPYCPYPPKPGIVPAKSCVGVCFPTSPVAVGT  
NTRSGAPTSWGENDTDVFLNNTPLGNWFGCTWMNSTGFTKVGAPCPVGGAG  
TLHCTDCFRKHDPATYSCSGSPWITPRCLVDYPYRLWHYPTCTNYTIFKRMVVG  
GYEHLREACNWRGRCDEDRSEIDMENTSGFLGLVLQAGFLLTLITIP  
QSLDSWNTSLNFGSPVCLGQNSQSPSTNSHSPCTPCPGYRWMLRRPIIFLIL  
LLCLIFLLVLLDYOQMLPVCPLIPGTTTSTGPKCTCTTPAQGNMFPSCCTKPTDG  
NCTCIPSPSWAFAPKYLMEWASVRFWSLSLLPVFQVFWGLSPTVLSAIWMWYKGP  
SLYSIVSPFIPILLPIFFCLMWYI"

ORIGIN  
Query Match 100.0%; Score 1593; DB 6; Length 5882;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTT 60  
 1992 ATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTT 2051  
 61 TGGCCAGCGCTAGCGAAACCCACAGTCACCGGGGAAAGTGC CGGCCACACACTGTGTCTGA 120  
 2052 TGGCCAGCGCTAGCGAAACCCACAGTCACCGGGGAAAGTGC CGGCCACACACTGTGTCTGA 2111  
 121 TTGTGTAGCTCTCTGCAACAGCGCCAAAGCAAGAGCTCCAGCTGATCAACACCGAC 180  
 2112 TTGTGTAGCTCTCTGCAACAGCGCCAAAGCAAGAGCTCCAGCTGATCAACACCGAC 2171  
 181 AGTTGGCACTCAATAGCAGCGCCCTGAAGCTGCAATGATAGCTCAACACCGCTGGTTG 240  
 2172 AGTTGGCACTCAATAGCAGCGCCCTGAAGCTGCAATGATAGCTCAACACCGCTGGTTG 2231  
 241 GCAGGGCTTTTCTATCACCAAGATTCAACTCTTCAAGGCTGTCTGAGAGGCTAGCCAGC 300  
 2232 GCAGGGCTTTTCTATCACCAAGATTCAACTCTTCAAGGCTGTCTGAGAGGCTAGCCAGC 2291  
 301 TGC CGA CCCCCTTACCGATTTTGA CAGGGCTGGGCCCTTATCAGTTATGCCAACGGAAGC 360  
 2292 TGC CGA CCCCCTTACCGATTTTGA CAGGGCTGGGCCCTTATCAGTTATGCCAACGGAAGC 2351  
 361 GGGCCGACAGCGCCCTACTGCTGGCACTACCCGCCAAACCTTGGCTATTTGGCC 420  
 2352 GGGCCGACAGCGCCCTACTGCTGGCACTACCCGCCAAACCTTGGCTATTTGGCC 2411  
 421 GGGAGAGTGTGTGTCGGTCCGGTATATTGCTTCACTCCAGCGCCCGTGGTGGGAACG 480  
 2412 GCGAGAGTGTGTGTCGGTCCGGTATATTGCTTCACTCCAGCGCCCGTGGTGGGAACG 2471  
 481 ACCGACAGTGTGGCGCGCCCACTACAGCTGGGGTGAAAAATGATACGGAAGTCTTCGTC 540  
 2472 ACCGACAGTGTGGCGCGCCCACTACAGCTGGGGTGAAAAATGATACGGAAGTCTTCGTC 2531  
 541 CTTTACAATACAGGCGCCGCTGGGCAATTGGTTCGGTTGTACCTGGATGAACCTCAACT 600  
 2532 CTTTACAATACAGGCGCCGCTGGGCAATTGGTTCGGTTGTACCTGGATGAACCTCAACT 2591  
 601 GGATTCAACCAAGTGTGGAGCGCCTCTTGTGTATCGGAGGGCGGGCAACACACC 660  
 2592 GGATTCAACCAAGTGTGGAGCGCCTCTTGTGTATCGGAGGGCGGGCAACACACC 2651  
 661 CTGCACTGCCCACTGATTGCTTCCGCAAGCAATCCGGAACGCACTCTCTGGTGGCG 720  
 2652 CTGCACTGCCCACTGATTGCTTCCGCAAGCAATCCGGAACGCACTCTCTGGTGGCG 2711  
 721 TCCGGTCCCTGGATCACCCAGGTGCTGCTGCACTACCGTATAGCGCTTTGGCATTAAT 780  
 2712 TCCGGTCCCTGGATCACCCAGGTGCTGCTGCACTACCGTATAGCGCTTTGGCATTAAT 2771  
 781 CTTTGTACCATCACTACACCATATTTAAATCAGGATGTACGTGGAGGGGTGCAACAC 840  
 2772 CTTTGTACCATCACTACACCATATTTAAATCAGGATGTACGTGGAGGGGTGCAACAC 2831  
 841 AGCTGGAAAGCTGCTGCAACTGGAACGCGGGCGAAAGCTTGGCATCTGGAAGATAGGAC 900  
 2832 AGCTGGAAAGCTGCTGCAACTGGAACGCGGGCGAAAGCTTGGCATCTGGAAGATAGGAC 2891  
 901 AGCTGGAGATCGATATGAGAACATCAATCAGGATTCCTAGGACCCCTGCTCGTTTA 960  
 2892 AGCTGGAGATCGATATGAGAACATCAATCAGGATTCCTAGGACCCCTGCTCGTTTA 2951  
 961 CAGCGCGGGTTTTTCTTGTGACAAAGATCCTCACAAATACCGCAGAGCTAGACTCGTGG 1020  
 2952 CAGCGCGGGTTTTTCTTGTGACAAAGATCCTCACAAATACCGCAGAGCTAGACTCGTGG 3011  
 1021 TGACCTTCTCAATTTTCTAGGGGATCTCCCGGTGTCTTGGCCAAAATTCGCAAGTCC 1080  
 3012 TGACCTTCTCAATTTTCTAGGGGATCTCCCGGTGTCTTGGCCAAAATTCGCAAGTCC 3071

1081 CCAACTCCAACTCACTCAACCAACCTCCTGTCTCCAAATTTGCTCGGTTATCGCTGATG 1140  
 3072 CCAACTCCAACTCACTCAACCAACCTCCTGTCTCCAAATTTGCTCGGTTATCGCTGATG 3131  
 1141 TGTCTGCGGCGTTTATCATATTCCTTTCATCTCTGCTGCTATGCTCATCTTCTTATTTG 1200  
 3132 TGTCTGCGGCGTTTATCATATTCCTTTCATCTCTGCTGCTATGCTCATCTTCTTATTTG 3191  
 1201 GTTCTTCTGGATTAACAAGTATGTGCGGTTGCTCTCTAAATTTCCAGGATCAACAACA 1260  
 3192 GTTCTTCTGGATTAACAAGTATGTGCGGTTGCTCTCTAAATTTCCAGGATCAACAACA 3251  
 1261 ACCAGTAGGGACCATGCAAAACCTGACGACTCTCTGCTCAAGGCAACTCTATGTTTCCC 1320  
 3252 ACCAGTAGGGACCATGCAAAACCTGACGACTCTCTGCTCAAGGCAACTCTATGTTTCCC 3311  
 1321 TCATGTTGCTGTACAAAACCTACGGATGGAATTCGACTGTTATTTCCCATCCATCCGTC 1380  
 3312 TCATGTTGCTGTACAAAACCTACGGATGGAATTCGACTGTTATTTCCCATCCATCCGTC 3371  
 1381 TGGGCTTTCCGCAAAATACCTATGGGAGTGGGCTCAGTCCGTTTCTTGGCTCAGTTTA 1440  
 3372 TGGGCTTTCCGCAAAATACCTATGGGAGTGGGCTCAGTCCGTTTCTTGGCTCAGTTTA 3431  
 1441 CTAGTGCCATTTGTTTCAGTGGTTCTAGGGCTTTCCGCTGTTTGGCTTTTCCAGCTATA 1500  
 3432 CTAGTGCCATTTGTTTCAGTGGTTCTAGGGCTTTCCGCTGTTTGGCTTTTCCAGCTATA 3491  
 1501 TGGATGATGTGGTATTTGGGGCCAAAGTCTGTACAGATCGTGAAGTCCCTTTATACCGCTG 1560  
 3492 TGGATGATGTGGTATTTGGGGCCAAAGTCTGTACAGATCGTGAAGTCCCTTTATACCGCTG 3551  
 1561 TTACCAATTTTCTTTTGTCTCTGCGTATACATT 1593  
 3552 TTACCAATTTTCTTTTGTCTCTGCGTATACATT 3584

BD132859 1989 bp DNA linear PAT 18-SEP-2002  
 Intracellular production of hepatitis C E1 and E2 truncated polypeptides.  
 BD132859.1 GI:23227804  
 JP 2002504810-A/2.  
 unclassified  
 unclassified  
 unclassified  
 1 (bases 1 to 1989)  
 Houghton, M., Choo, Q.L., Abrignani, S., Chien, D., Selby, M. and Glazer, E.  
 Intracellular production of hepatitis C E1 and E2 truncated Patent: JP 2002504810-A 2 12-FEB-2002;  
 CHIRON CORP  
 OS Hepatitis virus (hepatitis C virus)  
 PN JP 2002504810-A/2  
 PD 12-FEB-2002  
 PF 06-MAY-1998 JP 1998548360  
 PR 06-MAY-1997 US 60/045675  
 PI MICHAEL HOUGHTON, QUI LIM CHOO, SERGIO ABRIGNANI, DAVID CHIEN, MARK SELBY, EDWARD GLAZER  
 PC C12N15/40, C07K14/18, A61K39/29, G01N33/50  
 CC  
 FH Key Location/Qualifiers  
 FT CDS (1) . (1989) .  
 Location/Qualifiers  
 1. . 1989  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"

FEATURES  
 source  
 ORIGIN







Qy	507	CAGCTGGGGTGA	AATGATACCG	ACGTCTTCGTCTTAACAATACCAAGCCACC	GCTGGG	566
Db	565	CAGCTGGGGTGA	AATGATACCG	ACGTCTTCGTCTTAACAATACCAAGCCACC	GCTGGG	624
Qy	567	CAATTGGTTCCG	TGTACTCTCGATC	GNACCTAACTCGATTCACCAAAGTGCGGAGCGCC		625
Db	625	CAATTGGTTCCG	TGTACTCTCGATC	GNACCTAACTCGATTCACCAAAGTGCGGAGCGCC		684
Qy	627	TCCTTGTGTCA	TTCGAGGGGCGG	GCAACAACACCTCTGCATCGCCCCACTGATTGCTTCG		686
Db	685	TCCTTGTGTCA	TTCGAGGGGCGG	GCAACAACACCTCTGCATCGCCCCACTGATTGCTTCG		744
Qy	687	CAAGCATCCG	AGCCACATACTCTCGTGGGGTCCGGTCCCTGGATCACACCAAGTG			746
Db	745	CAAGCATCCG	AGCCACATACTCTCGTGGGGTCCGGTCCCTGGATCACACCAAGTG			804
Qy	747	CCTGGTCGACT	PACCCGCTATAGGCTTTGGCATTATTCCTTGTACCATCAACTACACCATATT			806
Db	805	CCTGGTCGACT	PACCCGCTATAGGCTTTGGCATTATTCCTTGTACCATCAACTACACCATATT			866
Qy	807	TAAAAATCAG	GATGTACGTGGAGGGGTCGAACACAGGCTGGAAGCTGCCTGCAACTGGAC			866
Db	865	TAAAAATCAG	GATGTACGTGGAGGGGTCGAACACAGGCTGGAAGCTGCCTGCAACTGGAC			924
Qy	867	GCGGGGCGA	CGTTGGCATCTTGGAAAGTAGGAGCAGGTCCGAGTC			912
Db	925	GCGGGGCGA	CGTTGGCATCTTGGAAAGTAGGAGCAGGTCCGAGTC			970
RESULT 6						
ARL18703						
LOCUS ARL18703 8316 bp DNA linear PAT 16-MAY-2001						
DEFINITION Sequence 88 from patent US 6150087.						
ACCESSION ARL18703						
VERSION ARL18703.1 GI:14100613						
KEYWORDS						
SOURCE Unknown.						
ORGANISM Unclassified.						
REFERENCE 1 (bases 1 to 8316)						
AUTHORS Chien,D.Y.						
TITLE NANBV diagnostics and vaccines						
JOURNAL Patent: US 6150087-A 88 21-NOV-2000;						
FEATURES Location/Qualifiers						
source						
1..8316						
/organism="unknown"						
/mol_type="unassigned DNA"						
 ORIGIN						
Query Match 52.4%; Score 834.8; DB 6; Length 8316;						
Best Local Similarity 96.4%; Pred. No. 3.3e-220;						
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;						
Qy	27	CTGTGTCTCTG	CTGTGTGGAGCAGTCTTCGTTTCGCCAGCGCTAGCGAAACCCACGT			86
Db	759	CTGGCGAAGGT	CTGTGTAGTGTCTGCTATTATTTGCCGGGCTGACCGGGAAACCCACGT			818
Qy	87	CACCGGGGAAG	TCCGGCCACACTGTGTCTGGATTTGTAGCTCTCTCGACACAGCGC			146
Db	819	CACCGGGGAAG	TCCGGCCACACTGTGTCTGGATTTGTAGCTCTCTCGACACAGCGC			878
Qy	147	CMAGCAAGCT	CCAGCTGTATCAACCAACGGCAGTTGGCACTCAATAGCAGCGCCT			206
Db	879	CAAGCAAG	CGTCCAGCTGTATCAACCAACGGCAGTTGGCACTCAATAGCAGCGCCT			938
Qy	207	GAACTGCAAT	ATAGCCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCAAGTT			266
Db	939	GAACTGCAAT	ATAGCCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCAAGTT			998
Qy	267	CAACTTTCAG	GCTGTCTGAGAGGCTAGCAGCTGCCGACCCCTTACCGATTTGACCA			326
Db	999	CAACTTTCAG	GCTGTCTGAGAGGCTAGCAGCTGCCGACCCCTTACCGATTTGACCA			1058

[illegible]

```
Qy 147 CAACGACGCTCAGCTGATCAACACACGCGCAGTTGGCACTCAATAGCAGCGCCT 206
Db 1540 CAACGACGCTCAGCTGATCAACACACGCGCAGTTGGCACTCAATAGCAGCGCCT 1599
Qy 207 GAACTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCAAAAGTT 266
Db 1600 GAACTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCAAAAGTT 1659
Qy 267 CAACCTCTTACGGCTGCTGAGAGGCTAGCAGCTGCGGACCCCTTACCGATTGACCA 326
Db 1660 CAACCTCTTACGGCTGCTGAGAGGCTAGCAGCTGCGGACCCCTTACCGATTGACCA 1719
Qy 327 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACAGCGCCCTACTGCTG 386
Db 1720 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACAGCGCCCTACTGCTG 1779
Qy 387 GCACCTACCCCCCAAAACCTTCGGGTATTGTGCCCGCGAAGAGTGTGTGTCGGGTATA 446
Db 1780 GCACCTACCCCCCAAAACCTTCGGGTATTGTGCCCGCGAAGAGTGTGTGTCGGGTATA 1839
Qy 447 TTGCTTCACTCCAGCCCGCTGGTGGGAACGACCGACAGGTGCGGCGGCCCACTTA 506
Db 1840 TTGCTTCACTCCAGCCCGCTGGTGGGAACGACCGACAGGTGCGGCGGCCCACTTA 1899
Qy 507 CAGCTGGGGTGAAATGATACGAGCTCTTCGTCCTTAACAAATACAGCGCCACCGCTGGG 566
Db 1900 CAGCTGGGGTGAAATGATACGAGCTCTTCGTCCTTAACAAATACAGCGCCACCGCTGGG 1959
Qy 567 CAATTGGTTTCGGTTGTAACCTGAACTCAACTGGAATTCACCAAGTGTGCGGAGCGCC 626
Db 1960 CAATTGGTTTCGGTTGTAACCTGGAATTCACCAAGTGTGCGGAGCGCC 2019
Qy 627 TCCTTGTGTCTATCGAGGGGGGGCAACAAACCCCTGCACTGCCCACTGATTGCTTCGG 686
Db 2020 TCCTTGTGTCTATCGAGGGGGGGCAACAAACCCCTGCACTGCCCACTGATTGCTTCGG 2079
Qy 687 CAAGCATCCGAGCGCCACATCTCTCGGTGGGCTCCGGTCCCTGGATCAACACCGAGTG 746
Db 2080 CAAGCATCCGAGCGCCACATCTCTCGGTGGGCTCCGGTCCCTGGATCAACACCGAGTG 2139
Qy 747 CCTGGTGCAGTACCCTGATAGGCTTTGGCATTTATCTTTGATCACTCAATCAACCATATT 806
Db 2140 CCTGGTGCAGTACCCTGATAGGCTTTGGCATTTATCTTTGATCACTCAATCAACCATATT 2199
Qy 807 TAAATATCAGGATGTACGTGGGAGGGGTGAAACACAGGCTGGAAGCTGCAACTGGAC 866
Db 2200 TAAATATCAGGATGTACGTGGGAGGGGTGAAACACAGGCTGGAAGCTGCAACTGGAC 2259
Qy 867 GCGGGCGGAACCTTGGATCTGGAAGATAGGACAGGTCGGAGATC 912
Db 2260 GCGGGCGGAACCTTGGATCTGGAAGATAGGACAGGTCGGAGATC 2305
```

```
RESULT 8
AR118723/c
LOCUS 118723 9185 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 123 from patent US 6150087.
ACCESSION AR118723
VERSION AR118723.1 GI:14100633
```

```
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 9185)
AUTHORS Chien,D.Y.
TITLE NANBV diagnostics and vaccines
JOURNAL Patent: US 6150087-A 123 21-NOV-2000;
FEATURES Location/Qualifiers
source 1. 9185
/mol_type="unassigned DNA"
```

ORIGIN

```
Query Match 52.4%; Score 834.8; DB 6; Length 9185;
Best Local Similarity 96.4%; Pred. No. 3.3e-220;
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
Qy 27 CTGTGTCTCTCTCTGTGTGGAGCAGCTCTTGGTTTGGCCCGAGCGCTAGCGAAGCCACAGT 86
Db 7766 CTGGCGGAAGTCTCTGTGTGTGTGTCTATTGGCCGGCGTGAACGGGAACCCACAGT 7707
Qy 87 CACCGGGGAAGTGGCGGCCACACTGTGTCTGGATTGTTAGCTCTCTCGACACAGCGC 146
Db 7706 CACCGGGGAAGTGGCGGCCACACTGTGTCTGGATTGTTAGCTCTCTCGACACAGCGC 7647
Qy 147 CAAGCAGAACGCTCCAGCTGATCAACACAAACGCGAGTTGGCACTCAATAGCAGCGCCCT 206
Db 7646 CAAGCAGAACGCTCCAGCTGATCAACACAAACGCGAGTTGGCACTCAATAGCAGCGCCCT 7587
Qy 207 GAACTGCAATGATAGCTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCAAAAGTT 266
Db 7586 GAACTGCAATGATAGCTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCAAAAGTT 7527
Qy 267 CAACCTCTTACGGCTGCTGAGAGGCTAGCAGCTGCGGACCCCTTACCGATTGACCA 326
Db 7526 CAACCTCTTACGGCTGCTGAGAGGCTAGCAGCTGCGGACCCCTTACCGATTGACCA 7467
Qy 327 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACAGCGCCCTACTGCTG 386
Db 7466 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACAGCGCCCTACTGCTG 7407
Qy 387 GCACCTACCCCCCAAAACCTTCGGGTATTGTGCCCGCGAAGAGTGTGTGTCGGGTATA 446
Db 7406 GCACCTACCCCCCAAAACCTTCGGGTATTGTGCCCGCGAAGAGTGTGTGTCGGGTATA 7347
Qy 447 TTGCTTCACTCCAGCCCGCTGGTGGGAACGACCGACAGGTGCGGCGGCCCACTTA 506
Db 7346 TTGCTTCACTCCAGCCCGCTGGTGGGAACGACCGACAGGTGCGGCGGCCCACTTA 7287
Qy 507 CAGCTGGGGTGAAATGATACGAGCTCTTCGTCCTTAACAAATACAGCGCCACCGCTGGG 566
Db 7286 CAGCTGGGGTGAAATGATACGAGCTCTTCGTCCTTAACAAATACAGCGCCACCGCTGGG 7227
Qy 567 CAATTGGTTTCGGTTGTAACCTGAACTCAACTGGAATTCACCAAGTGTGCGGAGCGCC 626
Db 7226 CAATTGGTTTCGGTTGTAACCTGGAATTCACCAAGTGTGCGGAGCGCC 7167
Qy 627 TCCTTGTGTCTATCGAGGGGGGGCAACAAACCCCTGCACTGCCCACTGATTGCTTCGG 686
Db 7166 TCCTTGTGTCTATCGAGGGGGGGCAACAAACCCCTGCACTGCCCACTGATTGCTTCGG 7107
Qy 687 CAAGCATCCGAGCGCCACATCTCTCGGTGGGCTCCGGTCCCTGGATCAACACCGAGTG 746
Db 7106 CAAGCATCCGAGCGCCACATCTCTCGGTGGGCTCCGGTCCCTGGATCAACACCGAGTG 7047
Qy 747 CCTGGTGCAGTACCCTGATAGGCTTTGGCATTTATCTTTGATCACTCAATCAACCATATT 806
Db 7046 CCTGGTGCAGTACCCTGATAGGCTTTGGCATTTATCTTTGATCACTCAATCAACCATATT 6987
Qy 807 TAAATATCAGGATGTACGTGGGAGGGGTGAAACACAGGCTGGAAGCTGCAACTGGAC 866
Db 6986 TAAATATCAGGATGTACGTGGGAGGGGTGAAACACAGGCTGGAAGCTGCAACTGGAC 6927
Qy 867 GCGGGCGGAACCTTGGATCTGGAAGATAGGACAGGTCGGAGATC 912
Db 6926 GCGGGCGGAACCTTGGATCTGGAAGATAGGACAGGTCGGAGATC 6881
```

```
RESULT 9
I08294
LOCUS I08294
DEFINITION Sequence 1 from Patent EP 0388232.
ACCESSION I08294
VERSION I08294.1 GI:588994
KEYWORDS linear PAT 02-DEC-1994
```

SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 9185)  
AUTHORS Houghton, M., Choo, Q.-L., and Kuo, G.  
TITLE NABV diagnostics and vaccines  
JOURNAL Patent: EP 0388232-A1 1 19-SEP-1990;  
FEATURES Location/Qualifiers  
source 1..9185  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 52.4%; Score 834.8; DB 6; Length 9185;  
Best Local Similarity 96.4%; Pred. No. 3.3e-220;  
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGCTCTGCTGTGTGAGAGTCTTCGTTTGGCCAGGCTAGCGAAACCCACGT 86  
DB 1420 CTGGCGAAGTCTCTGTAGTGTCTGCTATTTGGCGGCTCGACGCGGAAACCCACGT 1479

QY 87 CACCGGGGAAGTCCCGCCACACTGTCTGGATTGTAGCTCTCTGCACACGGCGC 146  
DB 1480 CACCGGGGAAGTCCCGCCACACTGTCTGGATTGTAGCTCTCTGCACACGGCGC 1539

QY 147 CAAGCAGAACGCTCCAGCTGATCAACCAACGCGAGTTGGCACCTCAATAGCACGGCCCT 206  
DB 1540 CAAGCAGAACGCTCCAGCTGATCAACCAACGCGAGTTGGCACCTCAATAGCACGGCCCT 1599

QY 207 GAACCTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCAAGTT 266  
DB 1600 GAACCTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCAAGTT 1659

QY 267 CAACTCTTCAGGCTGTCTGAGAGGTAGCCAGCTCCGACCCCTTACCGATTTCACCA 326  
DB 1660 CAACTCTTCAGGCTGTCTGAGAGGTAGCCAGCTCCGACCCCTTACCGATTTCACCA 1719

QY 327 GGGCTGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCGACGAGCGGCCCTTACTGCTG 386  
DB 1720 GGGCTGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCGACGAGCGGCCCTTACTGCTG 1779

QY 387 GCATCTACCCCAAAACCTTGGCGTATTTGGCCGGAAGAGTGTGTGTCGGTATA 446  
DB 1780 GCATCTACCCCAAAACCTTGGCGTATTTGGCCGGAAGAGTGTGTGTCGGTATA 1839

QY 447 TTCTTTCACTCCACAGCCCGTGTGTGGGAACGACGAGCTCCGGCGCGCCACCTA 506  
DB 1840 TTCTTTCACTCCACAGCCCGTGTGTGGGAACGACGAGCTCCGGCGCGCCACCTA 1899

QY 507 CAGCTGGGGTGAATATGATACGAGCTCTTCTGCTCTTAACTAACAGGCGCACCGCTGGG 566  
DB 1900 CAGCTGGGGTGAATATGATACGAGCTCTTCTGCTCTTAACTAACAGGCGCACCGCTGGG 1959

QY 567 CAATTGGTTGGTTGACCTGGATGAACCTCAACTGGAATCAACAAAGTGTGGAGCGGC 626  
DB 1960 CAATTGGTTGGTTGACCTGGATGAACCTCAACTGGAATCAACAAAGTGTGGAGCGGC 2019

QY 627 TCCTTGTGTCATCGGAGGGGGGCAACACACCTGCACTGCGCCCACTGATTGCTTCGG 686  
DB 2020 TCCTTGTGTCATCGGAGGGGGGCAACACACCTGCACTGCGCCCACTGATTGCTTCGG 2079

QY 687 CAAAGCATCCGGAGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 746  
DB 2080 CAAAGCATCCGGAGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 2139

QY 747 CTTGGTTCGACTACCCGTATAGGTTTGGCATTTATCTTGTACCATCAACTACACCATATT 806  
DB 2140 CTTGGTTCGACTACCCGTATAGGTTTGGCATTTATCTTGTACCATCAACTACACCATATT 2199

QY 807 TAAATTCAGGATGATCGTGGAGGGGTGCAACACAGGCTGGAGCTGCCTGCACTGGAC 866  
DB 2200 TAAATTCAGGATGATCGTGGAGGGGTGCAACACAGGCTGGAGCTGCCTGCACTGGAC 2259

QY 867 GCGGGGCGAACGTTGGGATCTGGAAGATAGGACAGGTCCGAGATC 912  
DB 2260 GCGGGGCGAACGTTGGGATCTGGAAGACAGGACAGGTCCGAGCTC 2305

RESULT 10  
BD091382 9185 bp DNA linear PAT 27-AUG-2002  
LOCUS HCV cultivation method in eucaryotic cells.  
DEFINITION  
ACCESSION BD091382  
VERSION BD091382.1 GI:22636993  
KEYWORDS JP 2001314192-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 9185)  
AUTHORS Weiner, A.J., Steimer, K.S. and Houghton, M.  
TITLE HCV cultivation method in eucaryotic cells  
JOURNAL Patent: JP 2001314192-A 3 13-NOV-2001;  
CHIRON CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2001314192-A/3  
PD 13-NOV-2001  
PF 15-MAR-2001 JP 2001075114  
PI 25-AUG-1989 US 398667  
PR AMY J WEINER, KATHELYN S STEIMER, MICHAEL HOUGHTON PC  
C12N15/09, C12N5/10, C12N7/00// (C12N7/00, C12R1.93), C12N15/00, PC  
C12N5/00  
CC HCV cultivation method in eucaryotic cells  
FH Key 1..9185  
FT source Location/Qualifiers  
FT /organism="Homo sapiens (human)".

FEATURES  
source 1..9185  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 52.4%; Score 834.8; DB 6; Length 9185;  
Best Local Similarity 96.4%; Pred. No. 3.3e-220;  
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGCTGCTGTGTGAGAGCTCTTCTGTTTCGCCAGCGCTAGCGAAACCCACGT 86  
DB 1420 CTGGCGAAGTCTCTGTAGTGTCTGCTATTTGGCGGCTCGACGCGGAAACCCACGT 1479

QY 87 CACCGGGGAAGTCCCGCCACACTGTGTCTGGATTGTAGCTCTCTGCACACGGCGC 146  
DB 1480 CACCGGGGAAGTCCCGCCACACTGTGTCTGGATTGTAGCTCTCTGCACACGGCGC 1539

QY 147 CAAGCAGAACGCTCCAGCTGATCAACCAACGCGAGTTGGCACCTCAATAGCACGGCCCT 206  
DB 1540 CAAGCAGAACGCTCCAGCTGATCAACCAACGCGAGTTGGCACCTCAATAGCACGGCCCT 1599

QY 207 GAACCTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCAAGTT 266  
DB 1600 GAACCTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCAAGTT 1659

QY 267 CAACTCTTCAGGCTGTCTGAGAGGTAGCCAGCTCCGACCCCTTACCGATTTCACCA 326  
DB 1660 CAACTCTTCAGGCTGTCTGAGAGGTAGCCAGCTCCGACCCCTTACCGATTTCACCA 1719

QY 327 GGGCTGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCGACGAGCGGCCCTTACTGCTG 386  
DB 1720 GGGCTGGGGCCCTATCAGTTATGCCAACGGAAGCGGCCGACGAGCGGCCCTTACTGCTG 1779

QY 387 GCATCTACCCCAAAACCTTGGCGTATTTGGCCGGAAGAGTGTGTGTCGGTATA 446  
DB 1780 GCATCTACCCCAAAACCTTGGCGTATTTGGCCGGAAGAGTGTGTGTCGGTATA 1839

QY 447 TTCTTTCACTCCACAGCCCGTGTGTGGGAACGACGAGCTCCGGCGCGCCACCTA 506  
DB 1840 TTCTTTCACTCCACAGCCCGTGTGTGGGAACGACGAGCTCCGGCGCGCCACCTA 1899

QY 507 CAGCTGGGGTGAATATGATACGAGCTCTTCTGCTCTTAACTAACAGGCGCACCGCTGGG 566  
DB 1900 CAGCTGGGGTGAATATGATACGAGCTCTTCTGCTCTTAACTAACAGGCGCACCGCTGGG 1959

QY 567 CAATTGGTTGGTTGACCTGGATGAACCTCAACTGGAATCAACAAAGTGTGGAGCGGC 626  
DB 1960 CAATTGGTTGGTTGACCTGGATGAACCTCAACTGGAATCAACAAAGTGTGGAGCGGC 2019

QY 627 TCCTTGTGTCATCGGAGGGGGGCAACACACCTGCACTGCGCCCACTGATTGCTTCGG 686  
DB 2020 TCCTTGTGTCATCGGAGGGGGGCAACACACCTGCACTGCGCCCACTGATTGCTTCGG 2079

QY 687 CAAAGCATCCGGAGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 746  
DB 2080 CAAAGCATCCGGAGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 2139

QY 747 CTTGGTTCGACTACCCGTATAGGTTTGGCATTTATCTTGTACCATCAACTACACCATATT 806  
DB 2140 CTTGGTTCGACTACCCGTATAGGTTTGGCATTTATCTTGTACCATCAACTACACCATATT 2199

QY 807 TAAATTCAGGATGATCGTGGAGGGGTGCAACACAGGCTGGAGCTGCCTGCACTGGAC 866  
DB 2200 TAAATTCAGGATGATCGTGGAGGGGTGCAACACAGGCTGGAGCTGCCTGCACTGGAC 2259

Db 1840 TTTGCTTCACTCCAGCCCCGTTGGTGGGAACGACCGACAGTGGGGCGCCACACCTA 1899  
QY 507 CAGCTGGGGTGAATAATGATACGGAAGTCTTCTGCTTAAACAATACAGGCCACCGCTGGG 566  
Db 1900 CAGCTGGGGTGAATAATGATACGGAAGTCTTCTGCTTAAACAATACAGGCCACCGCTGGG 1959  
QY 567 CAATTGGTTGGTTGCTGATGATGAACTCAACTGATGATTCACCAAGTGTGGGAGCGCC 626  
Db 1960 CAATTGGTTGGTTGCTGATGATGAACTCAACTGATGATTCACCAAGTGTGGGAGCGCC 2019  
QY 627 TCTTTGTGTCATCGGAGGGCGGCAACAACCTGCACTGCCCACTGATTGCTTCGG 686  
Db 2020 TCTTTGTGTCATCGGAGGGCGGCAACAACCTGCACTGCCCACTGATTGCTTCGG 2079  
QY 687 CAAGCATCCGAGCGCCACATACTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 746  
Db 2080 CAAGCATCCGAGCGCCACATACTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 2139  
QY 747 CTTGCTGCACTACCGCTATAGCTTTGGCAATTATCTTGTACCATCAACTACACCATATT 806  
Db 2140 CTTGCTGCACTACCGCTATAGCTTTGGCAATTATCTTGTACCATCAACTACACCATATT 2199  
QY 807 TAAATACAGATGTACGTGGAGGGGTGCAACACAGGCTGGAAGCTGCCTGCAACTGGAC 866  
Db 2200 TAAATACAGATGTACGTGGAGGGGTGCAACACAGGCTGGAAGCTGCCTGCAACTGGAC 2259  
QY 867 GCGGGCGCAAGCTTGGATCTCGAAGATAGGACAGGTCCGAGATC 912  
Db 2260 GCGGGCGCAAGCTTGGATCTCGAAGATAGGACAGGTCCGAGCTC 2305

RESULT 11  
LOCUS AR166930 9379 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 1 from patent US 6284249.  
ACCESSION AR166930  
VERSION AR166930.1 GI:16243325  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 9379)  
TITLES Barban,V.  
JOURNAL Fusion polypeptide having the C protein and E1 protein of hepatitis C virus  
PATENT: US 6284249-A 1 04-SEP-2001;  
FEATURES  
source Location/Qualifiers  
1..9379  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 52.4%; Score 834.8; DB 6; Length 9379;  
Best Local Similarity 96.4%; Pred. No. 3.3e-220;  
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
QY 27 CTGTGTGCTGCTGTGTGGAGCAGTCTTCTGTTTCGCCCGCAGCGCTAGCGAAACCCACGT 86  
Db 1420 CTGGGCGAAGTCTCTGTAGTGTCTGCTATTTCGCGGCTCGACGCGGAAACCCACGT 1479  
QY 87 CACCGGGGAAGTCCGCGGCACACTGTGTCTGGAATTTGTTAGCTCTCTCGCACCGGC 146  
Db 1480 CACCGGGGAAGTCCGCGGCACACTGTGTCTGGAATTTGTTAGCTCTCTCGCACCGGC 1539  
QY 147 CAAGCAGAAGCTCCAGCTGATCAACACAGGCGAGTTCGCACCTCAATAGCACGCGCT 206  
Db 1540 CAAGCAGAAGCTCCAGCTGATCAACACAGGCGAGTTCGCACCTCAATAGCACGCGCT 1599  
QY 207 GAACTGCAATGATGAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAGTT 266  
Db 1600 GAACTGCAATGATGAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAGTT 1659

QY 267 CAACCTTTTCAAGCTCTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTTTGACCA 326  
Db 1660 CAACCTTTTCAAGCTCTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTTTGACCA 1719  
QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGCGGGCCCCGACCAAGCGCCCTACTGTCTG 386  
Db 1720 GGGCTGGGGCCCTATCAGTTATGCCAAGCGGGCCCCGACCAAGCGCCCTACTGTCTG 1779  
QY 387 GCATACCCCCCAAAACCTTGGGTTATTTGTCGCCGGAAGAGTGTGTGTCGGGTATA 446  
Db 1780 GCATACCCCCCAAAACCTTGGGTTATTTGTCGCCGGAAGAGTGTGTGTCGGGTATA 1839  
QY 447 TTGCTTCACTCCAGCCCGCTGGTGGGAACAGACGACAGGTGCGGCGGCCACACCTA 506  
Db 1840 TTGCTTCACTCCAGCCCGCTGGTGGGAACAGACGACAGGTGCGGCGGCCACACCTA 1899  
QY 507 CAGCTGGGGTGAAATATGATACGGAGCTCTTCTGCTTAAACAATACAGGCCACCGCTGGG 566  
Db 1900 CAGCTGGGGTGAAATATGATACGGAGCTCTTCTGCTTAAACAATACAGGCCACCGCTGGG 1959  
QY 567 CAATTGGTTGGTTGCTGATGAACTCAACTGGATTTCACCAAGTGTGGGAGCGCC 626  
Db 1960 CAATTGGTTGGTTGCTGATGAACTCAACTGGATTTCACCAAGTGTGGGAGCGCC 2019  
QY 627 TCTTTGTGTCATCGGAGGGCGGCAACAACCTGCACTGCCCACTGATTGCTTCGG 686  
Db 2020 TCTTTGTGTCATCGGAGGGCGGCAACAACCTGCACTGCCCACTGATTGCTTCGG 2079  
QY 687 CAAGCATCCGAGCGCCACATACTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 746  
Db 2080 CAAGCATCCGAGCGCCACATACTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 2139  
QY 747 CTTGCTGCACTACCGCTATAGCTTTGGCAATTATCTTGTACCATCAACTACACCATATT 806  
Db 2140 CTTGCTGCACTACCGCTATAGCTTTGGCAATTATCTTGTACCATCAACTACACCATATT 2199  
QY 807 TAAATACAGATGTACGTGGAGGGGTGCAACACAGGCTGGAAGCTGCCTGCAACTGGAC 866  
Db 2200 TAAATACAGATGTACGTGGAGGGGTGCAACACAGGCTGGAAGCTGCCTGCAACTGGAC 2259  
QY 867 GCGGGCGCAAGCTTGGATCTCGAAGATAGGACAGGTCCGAGATC 912  
Db 2260 GCGGGCGCAAGCTTGGATCTCGAAGATAGGACAGGTCCGAGCTC 2305

RESULT 12  
LOCUS AR301300 9379 bp DNA linear PAT 12-JUN-2003  
DEFINITION Sequence 1 from patent US 6538123.  
ACCESSION AR301300  
VERSION AR301300.1 GI:31689076  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 9379)  
TITLES Barban,V.  
JOURNAL Vaccine composition for preventing or treating hepatitis C  
PATENT: US 6538123-A 1 25-MAR-2003;  
FEATURES  
source Location/Qualifiers  
1..9379  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 52.4%; Score 834.8; DB 6; Length 9379;  
Best Local Similarity 96.4%; Pred. No. 3.3e-220;  
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
QY 27 CTGTGTGCTGCTGTGTGGAGCAGTCTTCTGTTTCGCCCGCAGCGCTAGCGAAACCCACGT 86  
Db 1420 CTGGGCGAAGTCTCTGTAGTGTCTGCTATTTCGCGGCTCGACGCGGAAACCCACGT 1479

QY 87 CACCGGGGAAGTGC CGGCACACTGTGTCTGATTTGTTAGCTCTCGCACCAAGCGC 146  
Db CACCGGGGAAGTGC CGGCACACTGTGTCTGATTTGTTAGCTCTCGCACCAAGCGC 1539  
QY 147 CAAGCAGAGCTCCAGCTGATCAACCAACGAGCTGGCACCTCAATAGCAGGCGCT 206  
Db CAAGCAGAGCTCCAGCTGATCAACCAACGAGCTGGCACCTCAATAGCAGGCGCT 1599  
QY 207 GAATCTCAATGATAGCTCAACACCGGCTGGTGGCAGGCTTTTCTATCACCACAAGTT 266  
Db GAATCTCAATGATAGCTCAACACCGGCTGGTGGCAGGCTTTTCTATCACCACAAGTT 1659  
QY 267 CAATCTTCAAGGCTGTCTGAGAGCTAGCCAGCTGCCGACCTTACCGATTTGACCA 326  
Db CAATCTTCAAGGCTGTCTGAGAGCTAGCCAGCTGCCGACCTTACCGATTTGACCA 1719  
QY 327 GGGCTGGGCGCTTATCAGTTATGCCAACGGAAGCGGCCCGGACAGCGCCCTACTGCTG 386  
Db GGGCTGGGCGCTTATCAGTTATGCCAACGGAAGCGGCCCGGACAGCGCCCTACTGCTG 1779  
QY 387 GCATACCCCCCAAAACCTTGGCGTATTTGTCGGGGAAGAGTGTGTGTGGTCCGGTATA 446  
Db GCATACCCCCCAAAACCTTGGCGTATTTGTCGGGGAAGAGTGTGTGTGGTCCGGTATA 1839  
QY 447 TTGCTTCACTCCAGCCCCGTGTGTGGGAACGACCGACAGTGGCGGCGGCCACCTA 506  
Db TTGCTTCACTCCAGCCCCGTGTGTGGGAACGACCGACAGTGGCGGCGGCCACCTA 1899  
QY 507 CAGCTGGGTGAAAATGATACGAGCTCTTCGTCTTAAACATACCAAGCCACCGTGG 566  
Db CAGCTGGGTGAAAATGATACGAGCTCTTCGTCTTAAACATACCAAGCCACCGTGG 1959  
QY 567 CAATTTGGTTGGTTGACCTGGATGAATCAACTGATTTCAACAAAGTTGGGAGCGC 626  
Db CAATTTGGTTGGTTGACCTGGATGAATCAACTGATTTCAACAAAGTTGGGAGCGC 2019  
QY 627 TCTTTGTGTCATCGGAGGCGGCAACCAACCTGCACTGCCCACTGATTTGCTTCG 686  
Db TCTTTGTGTCATCGGAGGCGGCAACCAACCTGCACTGCCCACTGATTTGCTTCG 2079  
QY 687 CAAGCATCCGAGCGCCACATCTCTCGTGGGCTCGGTCCTGTGATCACACCAAGTG 746  
Db CAAGCATCCGAGCGCCACATCTCTCGTGGGCTCGGTCCTGTGATCACACCAAGTG 2139  
QY 747 CTTGTGCACTACCCGTATAGCTTTGGCATTTATCTTGTACCATCAACTACACCATATT 806  
Db CTTGTGCACTACCCGTATAGCTTTGGCATTTATCTTGTACCATCAACTACACCATATT 2199  
QY 807 TAAATCAGGATGTACGTGGAGGGTCAACACAGGCTGGAAGCTGCTGCAACTGGAC 866  
Db TAAATCAGGATGTACGTGGAGGGTCAACACAGGCTGGAAGCTGCTGCAACTGGAC 2259  
QY 867 GCGGGCGAAGCTTGGATCTCGAGATAGGACAGGTCGGAGATC 912  
Db GCGGGCGAAGCTTGGATCTCGAGATAGGACAGGTCGGAGATC 2305

RESULT 13  
LOCUS ARI176483 9401 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 9 from patent US 6312889.  
ACCESSION ARI176483  
VERSION ARI176483.1 GI:17918838  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9401)  
AUTHORS Houghton,M., Choo,Q.-L. and Kuo,G.  
TITLE Combinations of hepatitis c virus (HCV) antigens for use in  
immunassays for anti-HCV antibodies  
JOURNAL Patent: US 6312889-A 9 06-NOV-2001;  
FEATURES Location/Qualifiers

Source 1..9401  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 52.4%; Score 834.8; DB 6; Length 9401;  
Best Local Similarity 96.4%; Pred. No. 3.3e-220;  
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCAGCGCTAGCGAAACCCACG 86  
Db CTTGGCGGAAGTCTCTGTGTAGTCTGTCTATTTGCCGGGCTGCAGCGGAAACCCACG 1501  
QY 87 CACCGGGGAAGTGC CGGCACACTGTGTCTGATTTGTTAGCTCTCGCACCAAGCGC 146  
Db CACCGGGGAAGTGC CGGCACACTGTGTCTGATTTGTTAGCTCTCGCACCAAGCGC 1561  
QY 147 CAAGCAGAGCTCCAGCTGATCAACCAACGAGCTGGCACCTCAATAGCAGGCGCT 206  
Db CAAGCAGAGCTCCAGCTGATCAACCAACGAGCTGGCACCTCAATAGCAGGCGCT 1621  
QY 207 GAATCTCAATGATAGCTCAACACCGGCTGGTGGCAGGCTTTTCTATCACCACAAGTT 266  
Db GAATCTCAATGATAGCTCAACACCGGCTGGTGGCAGGCTTTTCTATCACCACAAGTT 1681  
QY 267 CAATCTTCAAGGCTGTCTGAGAGCTAGCCAGCTGCCGACCTTACCGATTTGACCA 326  
Db CAATCTTCAAGGCTGTCTGAGAGCTAGCCAGCTGCCGACCTTACCGATTTGACCA 1741  
QY 327 GGGCTGGGCGCTTATCAGTTATGCCAACGGAAGCGGCCCGGACAGCGCCCTACTGCTG 386  
Db GGGCTGGGCGCTTATCAGTTATGCCAACGGAAGCGGCCCGGACAGCGCCCTACTGCTG 1801  
QY 387 CAGCTACCCCCCAAAACCTTGGCGTATTTGTCGGGGAAGAGTGTGTGGTCCGGTATA 446  
Db GCACTACCCCCCAAAACCTTGGCGTATTTGTCGGGGAAGAGTGTGTGGTCCGGTATA 1861  
QY 447 TTGCTTCACTCCAGCCCCGTGTGTGGGAACGACCGACAGTGGCGGCGGCCACCTA 506  
Db TTGCTTCACTCCAGCCCCGTGTGTGGGAACGACCGACAGTGGCGGCGGCCACCTA 1921  
QY 507 CAGCTGGGTGAAAATGATACGAGCTCTTCGTCTTAAACATACCAAGCTCGGCTGG 566  
Db CAGCTGGGTGAAAATGATACGAGCTCTTCGTCTTAAACATACCAAGCTCGGCTGG 1981  
QY 567 CAATTTGGTTGGTTGACCTGGATGAATCAACTGATTTCAACAAAGTTGGGAGCGC 626  
Db CAATTTGGTTGGTTGACCTGGATGAATCAACTGATTTCAACAAAGTTGGGAGCGC 2041  
QY 627 TCTTTGTGTCATCGGAGGCGGCAACCAACCTGCACTGCCCACTGATTTGCTTCG 686  
Db TCTTTGTGTCATCGGAGGCGGCAACCAACCTGCACTGCCCACTGATTTGCTTCG 2101  
QY 687 CAAGCATCCGAGCGCCACATCTCTCGTGGGCTCGGTCCTGTGATCACACCAAGTG 746  
Db CAAGCATCCGAGCGCCACATCTCTCGTGGGCTCGGTCCTGTGATCACACCAAGTG 2161  
QY 747 CTTGTGCACTACCCGTATAGCTTTGGCATTTATCTTGTACCATCAACTACACCATATT 806  
Db CTTGTGCACTACCCGTATAGCTTTGGCATTTATCTTGTACCATCAACTACACCATATT 2221  
QY 807 TAAATCAGGATGTACGTGGAGGGTCAACACAGGCTGGAAGCTGCTGCAACTGGAC 866  
Db TAAATCAGGATGTACGTGGAGGGTCAACACAGGCTGGAAGCTGCTGCAACTGGAC 2281  
QY 867 GCGGGCGAAGCTTGGATCTCGAGATAGGACAGGTCGGAGATC 912  
Db GCGGGCGAAGCTTGGATCTCGAGATAGGACAGGTCGGAGATC 2327

RESULT 14  
LOCUS E66593 9401 bp DNA linear PAT 18-JUN-2001

—



Db	1742	GGGCTGGGGCCCTATCAGTTATGCAACGGAAGCGCCCGACAGCGCCCTACTGCTG	1801
Qy	387	GCACTACCCCCCAAAACCTTTGCGGTATTGTCGCGGAAGAGTGTGTGTCGGTATA	446
Db	1802	GCACTACCCCCCAAAACCTTTGCGGTATTGTCGCGGAAGAGTGTGTGTCGGTATA	1861
Qy	447	TTGCTTCACTCCAGCCCGTGTGGTGGGAAAGACCGACAGGTGCGGCGCGCCACCTA	506
Db	1862	TTGCTTCACTCCAGCCCGTGTGGTGGGAAAGACCGACAGGTGCGGCGCGCCACCTA	1921
Qy	507	CAGCTGGGTGAAATGATACGACGTCCTTCGTCCTTAACAATACAGGCCACGCTGGG	566
Db	1922	CAGCTGGGTGAAATGATACGACGTCCTTCGTCCTTAACAATACAGGCCACGCTGGG	1981
Qy	567	CAATTGGTTCGGTTGTACCTGGATGAATCAACTGGATTACCAAAAGTGTGGAGCGCC	626
Db	1982	CAATTGGTTCGGTTGTACCTGGATGAATCAACTGGATTACCAAAAGTGTGGAGCGCC	2041
Qy	627	TCCTTGTGTCATCGGAGGGCGGGCAACAACACCTGCACTGCCCTGCTTCCG	686
Db	2042	TCCTTGTGTCATCGGAGGGCGGGCAACAACACCTGCACTGCCCTGCTTCCG	2101
Qy	687	CAAGCATCCGGACGCCACATACCTCGGTGCGGCTCCGCTCCCTGGATCAACCCAGGTG	746
Db	2102	CAAGCATCCGGACGCCACATACCTCGGTGCGGCTCCGCTCCCTGGATCAACCCAGGTG	2161
Qy	747	CCTGGTCGACTACCCGATAGGCTTTGGCATTTACCTGTACCATCAACTACCATATT	806
Db	2162	CCTGGTCGACTACCCGATAGGCTTTGGCATTTACCTGTACCATCAACTACCATATT	2221
Qy	807	TAAATCAGGATGTACGTGGAGGGGTGGAACACAGGCTGGAAGCTGCCCTGCACTGGAC	866
Db	2222	TAAATCAGGATGTACGTGGAGGGGTGGAACACAGGCTGGAAGCTGCCCTGCACTGGAC	2281
Qy	867	GCGGGCGAAGCTTGGATCTGGAAGATAGGAGCAGGTCCGAGATC	912
Db	2282	GCGGGCGAAGCTTGGATCTGGAAGATAGGAGCAGGTCCGAGATC	2327

Search completed: April 15, 2005, 01:09:51  
Job time : 6941 secs



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 17:24:47 ; Search time 895 Seconds  
(without alignments)  
10536.478 Million cell updates/sec

Title: US-10-715-665-6\_COPY\_1992\_3584

Perfect score: 1593

Sequence: 1 atggatgcaatgaagagagg.....ttgtctctgggtatacatt 1593

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1593	100.0	5882	5 AAD06793	Aad06793 Plasmid p
2	836.4	52.5	1989	2 AAT18246	Aat18246 HCV E2/N
3	836.4	52.5	1989	2 AAV81371	AAV81371 Hepatitis
4	834.8	52.4	1207	2 AAQ40330	Aaq40330 Sequence
5	834.8	52.4	8316	2 AAQ05955	Aaq05955 Hepatitis
6	834.8	52.4	8316	3 AAQ75296	Aaa75296 cDNA sequ
7	834.8	52.4	8316	12 ADN35977	Adn35977 HCV cDNA
8	834.8	52.4	9185	2 AAQ05956	Aaq05956 Sense str
9	834.8	52.4	9185	2 AAQ10566	Aaq10566 Hepatitis
10	834.8	52.4	9185	2 AAX00459	Aax00459 Hepatitis
11	834.8	52.4	9185	2 AAX26737	Aax26737 Nucleotid
12	834.8	52.4	9185	2 ADF66068	Adf66068 Hepatitis
13	834.8	52.4	9185	3 AAA75297	Aaa75297 Sense str
14	834.8	52.4	9401	2 AAT12710	Aat12710 Hepatitis
15	834.8	52.4	9401	2 AAT99981	Aat99981 HCV polyp
16	834.8	52.4	9401	2 AAV09989	Aav09989 HCV polyp
17	834.8	52.4	9401	6 AAD35043	Aad35043 Hepatitis
18	834.8	52.4	9401	12 AD123106	Ad123106 Hepatitis
19	834.8	52.4	9401	13 ADR29358	Adr29358 Hepatitis
20	833.2	52.3	1914	8 AB270946	Abz70946 Hepatitis

21	833.2	52.3	1914	8	ACC71824	Acc71824 HCV-1 E1/
22	833.2	52.3	1914	12	ADO00769	Ado00769 HCV-1 E1/
23	833.2	52.3	9185	12	ADN35979	Adn35979 HCV cDNA
24	833.2	52.3	9379	2	AAQ36209	Aaq36209 Composite
25	833.2	52.3	9400	2	AAQ21744	Aaq21744 Compiled
26	818.8	51.4	9618	11	ADN33102	Adn33102 Hepatitis
27	811.8	51.0	9133	2	AAZ07656	Aaz07656 Nucleotid
28	748.4	47.0	1207	2	AAQ40331	Aaq40331 Sequence
29	745.2	46.8	9646	6	AAV59361	Aav59361 Hepatitis
30	745.2	46.8	9646	6	ABK87285	Abk87285 cDNA enco
31	745.2	46.8	9646	8	ACA62466	Aca62466 HCV H77 c
32	745.2	46.8	11674	3	AAZ36210	Aaz36210 Nucleotid
33	745.2	46.8	12980	2	AAV59364	Aav59364 Hepatitis
34	745.2	46.8	12980	6	ABK87286	Abk87286 Hepatitis
35	745.2	46.8	12980	8	ACA62469	Aca62469 DNA encod
36	745.2	46.8	13198	3	AAZ36211	Aaz36211 Nucleotid
37	745.2	46.8	16622	3	AAZ36212	Aaz36212 Nucleotid
38	742	46.6	1845	12	ADJ10455	Adj10455 DNA of th
39	742	46.6	2517	12	ADJ10457	Adj10457 DNA of th
40	742	46.6	9599	2	AAZ24832	Aaz24832 Infectiou
41	742	46.6	9599	4	AAF23491	Aaf23491 Infectiou
42	742	46.6	9599	4	AAC86938	Aac86938 Nucleotid
43	742	46.6	9599	12	ADJ56743	Adj56743 Hepatitis
44	742	46.6	9599	12	ADJ64255	Adj64255 Hepatitis
45	742	46.6	10803	10	ADD67945	Add67945 Modified

## ALIGNMENTS

### RESULT 1

AAD06793  
ID AAD06793 standard; DNA; 5882 BP.

XX AAD06793;

XX 06-AUG-2001 (first entry)

XX Plasmid pCMV-II-E2661-Sag.

XX Virus-like particle; immunogen; hepatitis B virus surface antigen; HBsAg;  
XX HBV; chimeric antigen; hepatitis C virus; HCV; virucide; vaccine;  
XX plasmid pCMV-II-E2661-SAG; cyclic; circular; ds.

XX Human cytomegalovirus.

XX Homo sapiens.

XX Hepatitis C virus.

XX Hepatitis B virus.

XX Bos sp.

XX Unidentified.

XX Chimeric.

XX Key Location/Qualifiers

FT CDS 1992..3587

FT /tag= a

FT /product= "HCV 661 E2 envelope glycoprotein-HBsAg S

FT domain"

FT misc\_feature 1992..2900

FT /tag= b

FT /note= "661 E2 coding sequence from HCV"

FT misc\_feature 2907..3587

FT /tag= c

FT /note= "HBsAg S domain coding sequence from HBV"

XX WO200138358-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032249.

XX 24-NOV-1999; 99US-0167224P.

XX (CHIR ) CHIRON CORP.



XX AAT18246;  
 AC 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 04-JUN-1996 (first entry)  
 XX HCV1 E2/NS2 region DNA.  
 DE HCV1: E1 protein; E2 protein; truncation; vaccine; diagnosis; therapy;  
 KW protein secretion; ds.  
 KW Hepatitis C virus.  
 OS  
 XX Key Location/Qualifiers  
 FH sig\_peptide 1..63  
 FT /\*tag= a  
 FT mat\_peptide 64..1986  
 FT /\*tag= b  
 XX WO9604301-A2.  
 XX 15-FEB-1996.  
 PD 26-JUL-1995; 95WO-US010035.  
 XX 29-JUL-1994; 94US-00282959.  
 PR 25-JUL-1995; 95US-00506608.  
 XX (CHIR ) CHIRON CORP.  
 XX Selby M, Houghton M;  
 PI WPI: 1996-129331/13.  
 DR P-PSDB; AAR92935.  
 XX Secretable hepatitis C virus E1 and E2 polypeptide(s) lacking all/part of  
 PT the membrane spanning domain - useful in vaccines, and for diagnostic and  
 PT therapeutic purposes, e.g. in assays for HCV.  
 XX Disclosure; Fig 4a-4c; 46pp; English.  
 XX The hepatitis c virus (HCV) E2/NS2 region (AAT18245) codes for the  
 CC envelope E2+NS2 proteins (AAR92935). Elimination of sequences encoding  
 CC the C-terminal membrane-spanning domain (approx. amino acids 337-361) of  
 CC E2 provides a truncated protein that can be secreted when produced  
 CC recombinantly in host cells. Truncated versions of the E1 protein (see  
 CC AAR92934) are similarly produced. When co-expressed or combined after  
 CC separate expression, the truncated E2 and E1 proteins form a complex. The  
 CC truncated proteins and complex are useful in vaccines and for diagnostic  
 CC and therapeutic applns. (Updated on 25-MAR-2003 to correct PR field.)  
 CC (Updated on 27-AUG-2003 to correct OS field.)  
 CC  
 SQ Sequence 1989 BP; 373 A; 586 C; 568 G; 462 T; 0 U; 0 Other;  
 Query Match 52.5%; Score 836.4; DB 2; Length 1989;  
 Best Local Similarity 96.5%; Pred. No. 2.4e-133;  
 Matches 855; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
 QY 27 CTGTGTGCTGCTGTGTGGAGAGTCTTTCGTTTCGCCAGCGCTAGCGAAACCCACGT 86  
 DB 12 CTGGCGAAGTCTGTGTAGTGTCTGTCTATTTTGGCGGCTCGACGCGAAACCCACGT 71  
 QY 87 CACCGGGGAGTGGCGGCACACTGTCTGTGATTTGTAGCTCTCTCGCACAGGCGC 146  
 DB 72 CACCGGGGAGTGGCGGCACACTGTCTGTGATTTGTAGCTCTCTCGCACAGGCGC 131  
 QY 147 CAAGCAGACGCTCCAGCTGTATCAACACACGCGAGTTGGCAGCTCAATAGCAGCGCCCT 206  
 DB 132 CAAGCAGACGCTCCAGCTGTATCAACACACGCGAGTTGGCAGCTCAATAGCAGCGCCCT 191  
 QY 207 GAATCTCAATGATAGCTCAACACCGGCTGGTTGGCAGGCGCTTTTATCACCACAAGTT 266

192 GAACTGCAATGATAGCTCAACACCGGCTGTGTTGGCAGGCGCTTTTATCACCACAAGTT 251  
 QY 267 CAACTCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCTTACCGATTTTGACCA 326  
 DB 252 CAACTCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCCGACCTTACCGATTTTGACCA 311  
 QY 327 GGGCTGGGCGCTTATCAGTTATGCCAAAGGAAAGCGGCCCGACGAGCGCCCTACTCTG 386  
 DB 312 GGGCTGGGCGCTTATCAGTTATGCCAAAGGAAAGCGGCCCGACGAGCGCCCTACTCTG 371  
 QY 387 GCACTACCCCGCCAAAACCTTGGGTTATGTCGCCGGAAGAGTGTGTGGTTCGGTATA 446  
 DB 372 GCACTACCCCGCCAAAACCTTGGGTTATGTCGCCGGAAGAGTGTGTGGTTCGGTATA 431  
 QY 447 TTGCTTCACTCCAGCGCCCGTGTGGGAAAGCGAGCAGCTCGGCGCGCCACCTA 506  
 DB 432 TTGCTTCACTCCAGCGCCCGTGTGGGAAAGCGAGCAGCTCGGCGCGCCACCTA 491  
 QY 507 CAGCTGGGCTGAAAATGATACGAGCTTCTTCTGCTTAAACAATACCAAGGCGCCGCTGG 566  
 DB 492 CAGCTGGGCTGAAAATGATACGAGCTTCTTCTGCTTAAACAATACCAAGGCGCCGCTGG 551  
 QY 567 CAATTGCTGGTGTGTACCTGGATGAACCTCACTGATTCACCAAGTGTGGGAGCGCC 626  
 DB 552 CAATTGCTGGTGTGTACCTGGATGAACCTCACTGATTCACCAAGTGTGGGAGCGCC 611  
 QY 627 TCCTTGTGTATCGGAGGCGCGGCAACACACCTGCTGCTGCTGCTGCTGCTGCTGCT 686  
 DB 612 TCCTTGTGTATCGGAGGCGCGGCAACACACCTGCTGCTGCTGCTGCTGCTGCTGCT 671  
 QY 687 CAAGCATCCGAGCGCCACATATCTCTCGTGGGCTCGGTTCCCTGGATCACACCGAGGT 746  
 DB 672 CAAGCATCCGAGCGCCACATATCTCTCGTGGGCTCGGTTCCCTGGATCACACCGAGGT 731  
 QY 747 CTGTGTGACTACCTGCTGATGCTTGGCATTTATCTTGTATCATCACTACACCATATT 806  
 DB 732 CTGTGTGACTACCTGCTGATGCTTGGCATTTATCTTGTATCATCACTACACCATATT 791  
 QY 807 TAAATCAGGATGTAGCTGGGAGGCTCGAAACACAGCTGGAAGCTGCGTCAACTGGAC 866  
 DB 792 TAAATCAGGATGTAGCTGGGAGGCTCGAAACACAGCTGGAAGCTGCGTCAACTGGAC 851  
 QY 867 GCGGGGCGAAGCTTGGATCTGGAAGATAGGAGCAGGTCCGAGATC 912  
 DB 852 GCGGGGCGAAGCTTGGATCTGGAAGATAGGAGCAGGTCCGAGATC 897  
 RESULT 3  
 AAV81371  
 ID AAV81371 standard; DNA; 1989 BP.  
 XX  
 AC AAV81371;  
 XX  
 DT 02-MAR-1999 (first entry)  
 XX  
 DE Hepatitis C virus E2 coding region.  
 KW Hepatitis C virus; E1; HCV; truncation; variant; PCR; amplification;  
 KW affinity chromatography; Galanthus nivalis; agglutinin; medicament;  
 XX detection; infection; ss.  
 OS Hepatitis C virus.  
 PH Key Location/Qualifiers  
 FT CDS 1..1989  
 FT /\*tag= a  
 FT /product= "E2 protein"  
 XX WO9850556-A2.  
 PN  
 XX 12-NOV-1998.  
 PD  
 XX 06-MAY-1998; 98WO-US009097.  
 PF

XX 06-MAY-1997; 97US-0045675P.  
 PR (CHIR ) CHIRON CORP.  
 XX  
 XX Houghton M, Choo Q, Abrignani S, Chien D, Selby M, Glazer E;  
 XX  
 XX WPI; 1999-034724/03.  
 DR P-PSDB; AAW67615.  
 XX  
 XX Methods for isolating truncated HCV E1 and E2 polypeptides - used in,  
 PT e.g. immunodiagnostic kits for diagnosis of HCV infection.  
 PT  
 XX Disclosure; Fig 2A-E; 65pp; English.  
 XX  
 CC This sequence represents the coding region for the Hepatitis C virus  
 CC (HCV) E2 protein. The invention relates to the intracellular production  
 CC and isolation of C-terminally truncated variants of the E2 protein. The  
 CC truncations start from about amino acid residue 500 of the E2 protein and  
 CC are generated by PCR amplification of the gene sequence. The truncated  
 CC proteins are then produced intracellularly in host cells which are  
 CC disrupted. The HCV proteins are then purified preferably by affinity  
 CC chromatography, especially using Galanthus nivalis agglutinin resin. The  
 CC HCV E1 and HCV E2 polypeptides can be used to manufacture a medicament  
 CC useful for detecting the presence or absence of HCV infection in an  
 CC individual. They can also be used in a immunodiagnostic test kit for  
 CC detecting HCV infection  
 XX  
 XX Sequence 1989 BP; 372 A; 586 C; 569 G; 462 T; 0 U; 0 Other;  
 SQ

Query Match 52.5%; Score 836.4; DB 2; Length 1989;  
 Best Local Similarity 96.5%; Pred. No. 2.4e-133;  
 Matches 855; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 27 CTGTGTGCTGCTGTGTGAGCAGCTTCTGTTTCGCCCGCAGCGCTAGCGAAACCCACG 86  
 DB 12 CTGGCGAGGTCCTGTAGTGTCTGTCTATTTGCGGGGTGCGCGGAACCCACCT 71  
 QY 87 CACCGGGGAAGTGC CGGCGCACACTGTGTCTGTGATTTGTTAGCTCTCGCACCGGCC 146  
 DB 72 CACCGGGGAAGTGC CGGCGCACACTGTGTCTGTGATTTGTTAGCTCTCGCACCGGCC 131  
 QY 147 CAAGCAGACGCTCCAGCTCATCAACACCAACGCGCAGTGCACCTCAATAGACGSCCT 206  
 DB 132 CAAGCAGACGCTCCAGCTCATCAACACCAACGCGCAGTGTGCACCTCAATAGACGSCCT 191  
 QY 207 GAACTGCAATGATAGCTCAACACCGGCTGTTGCGAGGGCTTTTCTATCACCACAAGTT 266  
 DB 192 GAACTGCAATGATAGCTCAACACCGGCTGTTGCGAGGGCTTTTCTATCACCACAAGTT 251  
 QY 267 CAACTCTTCAAGGCTGTCTGAGAGGCTAGCCAGTCCGACCCCTTACCGATTGTGACCA 326  
 DB 252 CAACTCTTCAAGGCTGTCTGAGAGGCTAGCCAGTCCGACCCCTTACCGATTGTGACCA 311  
 QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGCGAGCGGCCCGGACCGGCCCTACTGCTG 386  
 DB 312 GGGCTGGGGCCCTATCAGTTATGCCAAGCGAGCGGCCCGGACCGGCCCTACTGCTG 371  
 QY 387 GCACTACCCCGGCAAAACCTTGCCTGATTTGTCGCCGGAAGAGTGTGTGTGTCGGTATA 446  
 DB 372 GCACTACCCCGGCAAAACCTTGCCTGATTTGTCGCCGGAAGAGTGTGTGTGTCGGTATA 431  
 QY 447 TTGCTTCACTCCAGCCCGCTGTGTGGGAACGACCGA CAGTCCGGCGCGCCACCTTA 506  
 DB 432 TTGCTTCACTCCAGCCCGCTGTGTGGGAACGACCGA CAGTCCGGCGCGCCACCTTA 491  
 QY 507 CAGCTGGGTGAAATGATACGACCTCTTCGTCCTTACATACACGAGCCACCGCTGG 566  
 DB 492 CAGCTGGGTGAAATGATACGACCTCTTCGTCCTTACATACACGAGCCACCGCTGG 551  
 QY 567 CAAATTGGTTGGTTGTACCTGGATGAACCTCAACTGGATTCAACAAAGTGTGGGAGCGCC 626  
 DB 552 CAAATTGGTTGGTTGTACCTGGATGAACCTCAACTGGATTCAACAAAGTGTGGGAGCGCC 611

QY 627 TCCTTTGTGTCTCGAGGGCGGCAACAAACACCTGCACTGCCCACTGATTGCTTCG 686  
 DB 612 TCCTTTGTGTCTCGAGGGCGGCAACAAACACCTGCACTGCCCACTGATTGCTTCG 671  
 QY 687 CAAGCATCCGAGCGCCACATACCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 746  
 DB 672 CAAGCATCCGAGCGCCACATACCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 731  
 QY 747 CTGTGTGACATACCCGTATAGCTTTGGCATTTATCTTGTACCATCAACTACCATATT 806  
 DB 732 CTGTGTGACATACCCGTATAGCTTTGGCATTTATCTTGTACCATCAACTACCATATT 791  
 QY 807 TAAATTCAGGATGATAGCTGGGAGGGTGCMAACACAGGTGGAGCTGCTGCAACTGGAC 866  
 DB 792 TAAATTCAGGATGATAGCTGGGAGGGTGCMAACACAGGTGGAGCTGCTGCAACTGGAC 851  
 QY 867 GGGGGCGCAACCTTGCATCTGGAAGATAGGACAGGTCCGAGATC 912  
 DB 852 GCGGGCGCAACCTTGCATCTGGAAGATAGGACAGGTCCGAGCTC 897

RESULT 4  
 AAQ40330  
 ID AAQ40330 standard; cDNA; 1207 BP.  
 XX  
 AC AAQ40330;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 09-AUG-1993 (first entry)  
 XX  
 DE Sequence encoding glycoprotein E2/NS1 in clone HCV1.  
 XX  
 KW Hepatitis C virus; envelope protein; glycoprotein; E2/NS1;  
 KW diagnostic reagent; ss.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 2..1207  
 FT /\*tag= a  
 XX  
 PN EP537626-A1.  
 XX  
 PD 21-APR-1993.  
 XX  
 PF 08-OCT-1992; 92EP-00117191.  
 XX  
 PR 08-OCT-1991; 91JP-00260824.  
 XX  
 PA (NAHE-) NAT INST OF HEALTH.  
 XX  
 PI Miyamura T, Saito I, Harada S, Honda Y;  
 XX  
 DR WPI: 1993-127516/16.  
 DR P-PSDB; AAR34438.  
 XX  
 PT Diagnostic reagent for hepatitis C virus - comprises second envelope  
 PT protein or first non-structural protein encoded by HCV gene and has sugar  
 PT chain.  
 XX  
 PS Claim 8; Page 30-32; 58pp; English.  
 XX  
 CC Glycoprotein E2/NS1 is derived from the second envelope protein or first  
 CC non-structural protein encoded by the genome of HCV. The nucleic acid is  
 CC extracted from the serum of the patient of Hepatitis C. The serum is  
 CC pref. mixed with transfer RNA (tRNA) as a carrier of virus RNA. As a  
 CC technique of cloning cDNA from the nucleic acid, it is preferred to use  
 CC polymerase chain reaction method. In the reaction, any commercially  
 CC available random primers or synthesized DNA having a base sequence  
 CC similar to that of primer AS1 may be used as a primer. Representative  
 CC examples of sense primers include S1. (Updated on 25-MAR-2003 to correct  
 CC PN field.)

XX SQ Sequence 1207 BP; 241 A; 369 C; 336 G; 261 T; 0 U; 0 Other;  
 Query Match 52.4%; Score 834.8; DB 2; Length 1207;  
 Best Local Similarity 96.4%; Pred. No. 4.9e-133;  
 Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
 QY 27 CTGTGTGCTGCTGCTGTGGAGCAGTCTTTCGCTTTCGCCCGAGCGCTAGCGAAACCCACGT 86  
 DB 85 CTGGCGAAGGCTCTGCTAGTGTCTGTCTATTTGCGCGCTGACGCGAAACCCACGT 144  
 QY 87 CACCGGGGAAGTCCCGGCCACACTGTGTCTGGAATTTGTAGCTCTCGCACCGGC 146  
 DB 145 CACCGGGGAAGTCCCGGCCACACTGTGTCTGGAATTTGTAGCTCTCGCACCGGC 204  
 QY 147 CAAGCAGAACGCTCCAGCTGATCAACACCAACGACGCTGCGACCTCAATAGCAGCGCCCT 206  
 DB 205 CAAGCAGAACGCTCCAGCTGATCAACACCAACGACGCTGCGACCTCAATAGCAGCGCCCT 264  
 QY 207 GAACCTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCACAAGTT 266  
 DB 265 GAACCTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCACAAGTT 324  
 QY 267 CAACCTTCAGGCTGCTGAGAGGCTAGCAGCTGCCAGCCCTTACCGATTTTGACCA 326  
 DB 325 CAACCTTCAGGCTGCTGAGAGGCTAGCAGCTGCCAGCCCTTACCGATTTTGACCA 384  
 QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACGAGCGCCCTACTGCTG 386  
 DB 385 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACGAGCGCCCTACTGCTG 444  
 QY 387 GCACCTACCCCAAAACCTTGGGTATTTGCCCGGAGAGTGTGTGTCGCGTATA 446  
 DB 445 GCACCTACCCCAAAACCTTGGGTATTTGCCCGGAGAGTGTGTGTCGCGTATA 504  
 QY 447 TTGCTTCACTCCAGCCCGTGTGTGGGAAGCAGCAGGTGTGGGGCGGCCACCTA 506  
 DB 505 TTGCTTCACTCCAGCCCGTGTGTGGGAAGCAGCAGGTGTGGGGCGGCCACCTA 564  
 QY 507 CAGCTGGGGTGAATATGATACGAGCTCTTCGCTTACCAATACACGACCAACCGCTGG 566  
 DB 565 CAGCTGGGGTGAATATGATACGAGCTCTTCGCTTACCAATACACGACCAACCGCTGG 624  
 QY 567 CAATTGGTTGCTGTGCTGAGTCACTCACTGATTTACCAAAAGTGTGCGAGCGCC 626  
 DB 625 CAATTGGTTGCTGTGCTGAGTCACTCACTGATTTACCAAAAGTGTGCGAGCGCC 684  
 QY 627 TCCTTGTGTCATCGAGGGCGGGCAACACACCTGCACTGCCCGCAGTGTCTCG 686  
 DB 685 TCCTTGTGTCATCGAGGGCGGGCAACACACCTGCACTGCCCGCAGTGTCTCG 744  
 QY 687 CAAGCATCCGAGCGGCACATCTCGGTGCGGCTCCGTCCTCGATCACACCCAGGTG 746  
 DB 745 CAAGCATCCGAGCGGCACATCTCGGTGCGGCTCCGTCCTCGATCACACCCAGGTG 804  
 QY 747 CTGTGTCGACTACCGGTATAGGCTTTGGCATTTATCTTGTACCATCAATACACCATATT 806  
 DB 805 CTGTGTCGACTACCGGTATAGGCTTTGGCATTTATCTTGTACCATCAATACACCATATT 864  
 QY 807 TAAATATCAGGATGATCGTGGAGGGGTGAAACACAGGCTGGAAGTGTGCTGCAATGAC 866  
 DB 865 TAAATATCAGGATGATCGTGGAGGGGTGAAACACAGGCTGGAAGTGTGCTGCAATGAC 924  
 QY 867 GCGGGGCGAAGCTTCGGATCTGGAAGATAGGACAGGTCGGATC 912  
 DB 925 GCGGGGCGAAGCTTCGGATCTGGAAGATAGGACAGGTCGGATC 970

RESULT 5  
 AAQ05955  
 ID AAQ05955 standard; cDNA; 8316 BP.  
 XX  
 AC AAQ05955;

XX 25-MAR-2003 (revised)  
 DT 23-JAN-1991 (first entry)  
 XX Hepatitis C virus open reading frame.  
 DE Hepatitis C virus; antiviral agent; ds.  
 KW Hepatitis C virus.  
 OS Hepatitis C virus.  
 XX EP388232-A.  
 XX 19-SEP-1990.  
 XX 16-MAR-1990; 90EP-00302866.  
 XX 17-MAR-1989; 89US-00325338.  
 PR 20-APR-1989; 89US-00341334.  
 PR 18-MAY-1989; 89US-00355002.  
 XX (CHIR ) CHIRON CORP.  
 PA Houghton M, Choo QL, Kuo G;  
 PI WPI; 1990-284418/38.  
 DR P-PSDB; AAR08123.  
 XX Hepatitis C virus DNA - used for producing probes, polypeptide(s),  
 PT antibodies and anti-sense polynucleotide(s) for diagnosis and therapy.  
 PT Disclosure; Fig 16; 83pp; English.  
 PS HCV cDNA libraries were constructed using pooled serum from a chimpanzee  
 CC with chronic HCV infection. A lambda gt10 library was screened with  
 CC probes derived from previously isolated clones. The ORF is derived from  
 CC the overlapping clones p14a, CA167b, CA156e, CA84a, CA59a, K9-1, 12f,  
 CC 14i, 11b, 7f, 8h, 33c, 40b, 37b, 36, 81, 32, 33b, 25c, 14c, 8f, 33f,  
 CC 33g, 35f, 19g, 26g and 15e. Polypeptide encoded by this sequence can  
 CC be used to design probes for the detection of HCV nucleic acids, in  
 CC screening programmes for antiviral agents and in preparing blood free of  
 CC HCV. Antisense polynucleotides can be used to inhibit viral replication.  
 CC See also AAQ05956. (Updated on 25-MAR-2003 to correct PA field.)  
 XX SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;  
 Query Match 52.4%; Score 834.8; DB 2; Length 8316;  
 Best Local Similarity 96.4%; Pred. No. 3.6e-133;  
 Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
 QY 27 CTGTGTGCTGCTGCTGTGGAGCAGTCTTTCGCTTTCGCCCGAGCGCTAGCGAAACCCACGT 86  
 DB 759 CTGGCGAAGTCTCTGTGTAGTGTCTCTATTTGCCGCGCTGAGCGGAACCCACGT 818  
 QY 87 CACCGGGGAAGTCCCGGCCACACTGTGTCTGGAATTTGTAGCTCTCGCACCGGC 146  
 DB 819 CACCGGGGAAGTCCCGGCCACACTGTGTCTGGAATTTGTAGCTCTCGCACCGGC 878  
 QY 147 CAAGCAGAACGCTCCAGCTGATCAACACCAACGACGCTGCGACCTCAATAGCAGCGCCCT 206  
 DB 879 CAAGCAGAACGCTCCAGCTGATCAACACCAACGACGCTGCGACCTCAATAGCAGCGCCCT 938  
 QY 207 GAACCTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCACAAGTT 266  
 DB 939 GAACCTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCACAAGTT 998  
 QY 267 CAACCTTCAGGCTGCTCTGAGAGGCTAGCAGCTGCCAGCCCTTACCGATTTTGACCA 326  
 DB 999 CAACCTTCAGGCTGCTCTGAGAGGCTAGCAGCTGCCAGCCCTTACCGATTTTGACCA 1058  
 QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGGGGCCCGACGAGCGCCCTACTGCTG 386  
 DB 1059 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGGGGCCCGACGAGCGCCCTACTGCTG 1118

QY 387 GCACTACCCCAAAACCTTCGGTATTGTGCCCGGAAGAGTGTGTGTGTCGGGTATA 446  
 DB 1119 GCACTACCCCAAAACCTTCGGTATTGTGCCCGGAAGAGTGTGTGTGTCGGGTATA 1178  
 QY 447 TTGCTTCACTCCAGCCCCCGTGGTGGGAACGACGACAGTGTGGGCGGCCACCTA 506  
 DB 1179 TTGCTTCACTCCAGCCCCCGTGGTGGGAACGACGACAGTGTGGGCGGCCACCTA 1238  
 QY 507 CAGCTGGGTGAAATGATACGAGCTCTTCGTCCTTAACAATACAGGCCACCGCTGGG 566  
 DB 1239 CAGCTGGGTGAAATGATACGAGCTCTTCGTCCTTAACAATACAGGCCACCGCTGGG 1298  
 QY 567 CAATTGGTTCGGTGTACTCGATGAACCTCAACTGAGTTTCCAAAGTGTGGGAGGCC 626  
 DB 1299 CAATTGGTTCGGTGTACTCGATGAACCTCAACTGAGTTTCCAAAGTGTGGGAGGCC 1358  
 QY 627 TCCTTGTGTCACTCGAGGGGGGCAACACACCTGCACTGCCCACTGATTGCTTCG 686  
 DB 1359 TCCTTGTGTCACTCGAGGGGGGCAACACACCTGCACTGCCCACTGATTGCTTCG 1418  
 QY 687 CAGCATCGGACGCCACATCTCTCGTGGGCTCCGTCCTGGATCACACCCAGGTG 746  
 DB 1419 CAGCATCGGACGCCACATCTCTCGTGGGCTCCGTCCTGGATCACACCCAGGTG 1478  
 QY 747 CCTGTGCACTACCCGTATAGGCTTTGGCATTTATCTTTGTACCATCAACTACCATATT 806  
 DB 1479 CCTGTGCACTACCCGTATAGGCTTTGGCATTTATCTTTGTACCATCAACTACCATATT 1538  
 QY 807 TAAATCAGGATGTACGTGGGAGGGTGAACACAGGCTGGAAGTGGCTGCACTGGAC 866  
 DB 1539 TAAATCAGGATGTACGTGGGAGGGTGAACACAGGCTGGAAGTGGCTGCACTGGAC 1598  
 QY 867 GCGGGCGAACGTTGGATCTGGAAGATAGGACAGGTCGAGATC 912  
 DB 1599 GCGGGCGAACGTTGGATCTGGAAGATAGGACAGGTCGAGATC 1644

RESULT 6

AAA75296  
 ID AAA75296 standard; cDNA; 8316 BP.

XX AAA75296;

XX 15-JAN-2001 (first entry)

XX cDNA sequence compiled Hepatitis C virus cDNA clones.

XX Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;  
 XX viral infectivity; viral replication; ds.

XX Hepatitis C virus.

XX Key Location/Qualifiers  
 FH 1..8316  
 FT /\*tag= a  
 FT /note= "partial sequence; no termination codon given"

XX EP1034785-A2.

XX 13-SEP-2000.

XX 16-MAR-1990; 2000EP-00109602.

XX 17-MAR-1989; 89US-00325338.

PR 20-APR-1989; 89US-00341334.

PR 18-MAY-1989; 89US-00355002.

PR 16-MAR-1990; 90EP-00302866.

XX (CHIR ) CHIRON CORP.

XX Houghton M, Choo Q, Kuo G;

XX WPI; 2000-566891/53.

DR P-PSDB; AAB18540.  
 XX Novel composition comprising a hepatitis C virus antisense polynucleotide  
 PT which is complementary to or corresponds to a sense strand of the virus  
 PT genome, and selectively hybridizes to it.  
 XX Example; Fig 16; 75pp; English.  
 XX The specification describes a pharmaceutical composition which comprises  
 CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is  
 CC characterized by a positive stranded RNA genome which has 40% homology at  
 CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide  
 CC binds to cellular polynucleotides which enhance and/or are required for  
 CC viral infectivity, replicative ability or chronicity. The antisense  
 CC polynucleotides may also be designed to bind with high specificity, to be  
 CC of increased stability, to be stable and to have low toxicity. The  
 CC composition also comprises an agent which causes viral RNA to be  
 CC inactive. The composition is used for preventing HCV replication in a  
 CC system. The present sequence represents a novel HCV cDNA sequence, which  
 CC is used in the course of the invention  
 SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;

Query Match 52.4%; Score 834.8; DB 3; Length 8316;  
 Best Local Similarity 96.4%; Pred. No. 3.6e-133;  
 Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGTGCTGCTGTGTGGAGCAGTCTTCGTTTGCCTCCAGCGCTAGCAAAACCCAGT 86  
 DB 759 CTGGCGGAGTCTGTGTAGTGTCTATTTTCGCGGCTCGACGGGAAACCCAGT 818  
 QY 87 CACCGGGGAAGTCCGGCCCACTGTCTGTGATTTGTAGCTCTTCGCACACGAGGC 146  
 DB 819 CACCGGGGAAGTCCGGCCCACTGTCTGTGATTTGTAGCTCTTCGCACACGAGGC 878  
 QY 147 CAACGAGAACGTCAGCTGATCAACCAACGAGTTGGCACTCAATAGCACGCCCT 206  
 DB 879 CAACGAGAACGTCAGCTGATCAACCAACGAGTTGGCACTCAATAGCACGCCCT 938  
 QY 207 GAACCTGCAATGATAGCTCAACACCGGCTGGTGTGACGGGCTTTTCTATCACCAAGTT 266  
 DB 939 GAACCTGCAATGATAGCTCAACACCGGCTGGTGTGACGGGCTTTTCTATCACCAAGTT 998  
 QY 267 CAACCTTTGAGGCTGTCTGAGAGCTAGCAGCTGCCGCCCTTACCGATTTGACCA 326  
 DB 999 CAACCTTTGAGGCTGTCTGAGAGCTAGCAGCTGCCGCCCTTACCGATTTGACCA 1058  
 QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGGGGGCCCGACAGCGCCCTACTGCTG 386  
 DB 1059 GGGCTGGGGCCCTATCAGTTATGCCAAGGGGGCCCGACAGCGCCCTACTGCTG 1118  
 QY 387 GCACTACCCCCCAAAACCTTCGGGTATTGTGCCCGCGAAGAGTGTGTGTGTCGGTATA 446  
 DB 1119 GCACTACCCCCCAAAACCTTCGGGTATTGTGCCCGCGAAGAGTGTGTGTGTCGGTATA 1178  
 QY 447 TTGCTTCACTCCAGCCCCCGTGGTGGGAACGACAGTGTGGGCGGCCACCTA 506  
 DB 1179 TTGCTTCACTCCAGCCCCCGTGGTGGGAACGACAGTGTGGGCGGCCACCTA 1238  
 QY 507 CAGCTGGGGTGAATGATACGAGCTCTTCGTCCTTAACAATACAGGCCACCGCTGGG 566  
 DB 1239 CAGCTGGGGTGAATGATACGAGCTCTTCGTCCTTAACAATACAGGCCACCGCTGGG 1298  
 QY 567 CAATTGGTTCGGTGTACTCGATGAACCTCAACTGGATTTCACCAAGTGTGGGAGGCC 626  
 DB 1299 CAATTGGTTCGGTGTACTCGATGAACCTCAACTGGATTTCACCAAGTGTGGGAGGCC 1358  
 QY 627 TCCTTGTGTCACTCGAGGGGGGCAACCAACCTTGCATGCCCACTGATTGCTTCG 686  
 DB 1359 TCCTTGTGTCACTCGAGGGGGGCAACCAACCTTGCATGCCCACTGATTGCTTCG 1418  
 QY 687 CAAGCATCCGAGCGGCACATCTCTCGGTGGGCTCCGGTCCCTGGATCACACCCAGGTG 746



Db 1419 CAAGCATCCGACGCCACATACCTCTCGGTGCGCTCCGGTCCCTGGATCACACCCAGGTG 1478  
QY 747 CTGTGTCGACTACCCGTATAGGCTTTGGCAATTATCTTTGTACCATCAACTACACCATATT 806  
Db 1479 CTGTGTCGACTACCCGTATAGGCTTTGGCAATTATCTTTGTACCATCAACTACACCATATT 1538  
QY 807 TAAATCAGGATGCTAGTGGAGGGTTCGAACAACAGGCTGGAGTGCCTGCAACTGGAC 866  
Db 1539 TAAATCAGGATGCTAGTGGAGGGTTCGAACAACAGGCTGGAGTGCCTGCAACTGGAC 1598  
QY 867 GCGGGCGGACGTTGCGATCTGGAAGTAGGACAGGTCCGAGATC 912  
Db 1599 GCGGGCGGACGTTGCGATCTGGAAGTAGGACAGGTCCGAGATC 1644

RESULT 7  
ADN35977  
ID ADN35977 standard; cDNA; 8316 BP.  
XX  
AC ADN35977;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
DE HCV cDNA clone #1.  
XX  
KW Antiviral; Vaccine; hepatitis C virus infection; HCV infection; ss.  
XX  
OS Hepatitis C virus.  
XX  
PN EP1394255-A2.  
XX  
PD 03-MAR-2004.  
XX  
PF 16-MAR-1990; 2003EP-00016585.  
XX  
PR 17-MAR-1989; 89US-00325338.  
PR 20-APR-1989; 89US-00341334.  
PR 18-MAY-1989; 89US-00355002.  
PR 16-MAR-1990; 90EP-00302866.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Houghton M, Choo Q, Kuo G;  
XX  
DR WPI; 2004-193149/19.  
DR P-PSDB; ADN35976.  
XX  
PT Novel purified hepatitis C virus polypeptide comprising epitope encoded  
PT by hepatitis C virus cDNA, useful as vaccine for treating hepatitis C  
PT virus.  
XX  
PS Example 1; Fig 16; 79pp; English.  
XX  
CC The present invention relates to hepatitis C virus (HCV) proteins and  
CC cDNA sequences. The sequences are useful in immunoassays for detecting  
CC antibodies directed against HCV antigen; preparing host cells transformed  
CC with a recombinant polynucleotide; screening antiviral agents and  
CC determining the effect of antiviral agent in inhibiting viral replication  
CC in cell culture system; and developing vaccine for treating HCV  
CC infection.  
XX  
SQ Sequence 8316 BP; 1671 A; 2537 C; 2338 G; 1770 T; 0 U; 0 Other;

Query Match 52.4%; Score 834.8; DB 12; Length 8316;  
Best Local Similarity 96.4%; Pred: No. 3.6e-133; Indels 0; Gaps 0;  
Matches 854; Conservative 0; Mismatches 32;  
QY 27 CTGTGTCGCTGCTGTGGAGCAGTCTTCGTTTCGCCGAGCGCTAGCGAAACCCACGT 86  
Db 759 CTGGGCGAAGTCTCTGGTAGTCTGCTGCTATTTCGCGGCTGAGCGGAAACCCACGT 818  
QY 87 CACCGGGGAAGTGGCGGCCACACTGTGTCTGGATTGTGCTCTCTCGACACCGAGGC 146

Db 819 CACCGGGGAAGTGGCGGCCACACTGTGTCTGGATTGTGTAGCTCTCTGCACACGCGC 878  
QY 147 CAAGCAGAAGCTCCAGCTGATCAACACCAACGCGAGTGTGCGACCTCAATAGACGCGCCT 206  
Db 879 CAAGCAGAAGCTCCAGCTGATCAACACCAACGCGAGTGTGCGACCTCAATAGACGCGCCT 938  
QY 207 GAATGCAATGATAGCTTCAACACACCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAGTT 266  
Db 939 GAATGCAATGATAGCTTCAACACACCGGCTGGTTGGCAGGGCTTTTCTATCACCACAAGTT 998  
QY 267 CAATCTTCAGGCTGTCTGAGAGGCTAGCAGCTCCGACCCCTTACCGATTTGACCA 326  
Db 999 CAATCTTCAGGCTGTCTGAGAGGCTAGCAGCTCCGACCCCTTACCGATTTGACCA 1058  
QY 327 GGGCTGGGCGCTTATCAGTTATGCAACGGAAGCGGCCCGACGCGCCCTACTGCTG 386  
Db 1059 GGGCTGGGCGCTTATCAGTTATGCAACGGAAGCGGCCCGACGCGCCCTACTGCTG 1118  
QY 387 GCACCTACCCCAAAACCTTTCGGTATTGTGCCGCGAAGAGTGTGTGTGTCGGTATA 446  
Db 1119 GCACCTACCCCAAAACCTTTCGGTATTGTGCCGCGAAGAGTGTGTGTGTCGGTATA 1178  
QY 447 TTGCTTCACTCCAGCCCGCTGTGTGGAAAGCAGCAGGTCCGGCGCGCCACCTA 506  
Db 1179 TTGCTTCACTCCAGCCCGCTGTGTGGAAAGCAGCAGGTCCGGCGCGCCACCTA 1238  
QY 507 CAGCTGGGGTGAAATGATACGAGCTCTTCGTCCTTAACAATACAGGCGCCAGCTGGG 566  
Db 1239 CAGCTGGGGTGAAATGATACGAGCTCTTCGTCCTTAACAATACAGGCGCCAGCTGGG 1298  
QY 567 CAATTGGTTTCGGTTCGTACCTGGATGAATCAACTGGAATTCACCAAGTGTGCGAGCGCC 626  
Db 1299 CAATTGGTTTCGGTTCGTACCTGGATGAATCAACTGGAATTCACCAAGTGTGCGAGCGCC 1358  
QY 627 TCCTTGTGTCATCGAGGGCGGGGCAACAAACCTTGCATGCGCCCACTGATTGCTTCG 686  
Db 1359 TCCTTGTGTCATCGAGGGCGGGGCAACAAACCTTGCATGCGCCCACTGATTGCTTCG 1418  
QY 687 CAAGCATCCGAGCGCCACATCTCTCGGTGCGGCTCCGCTCCGTGATCACACCCAGGTG 746  
Db 1419 CAAGCATCCGAGCGCCACATCTCTCGGTGCGGCTCCGCTCCGTGATCACACCCAGGTG 1478  
QY 747 CTTGTCGACTACCCGTATAGCTTTTGGCATTTATCTTGTACCATCAACTACACCATATT 806  
Db 1479 CTTGTCGACTACCCGTATAGCTTTTGGCATTTATCTTGTACCATCAACTACACCATATT 1538  
QY 807 TAAATCAGGATGCTAGTGGAGGGTTCGAACAACAGGCTGGAGCTGCTGCAACTGGAC 866  
Db 1539 TAAATCAGGATGCTAGTGGAGGGTTCGAACAACAGGCTGGAGCTGCTGCAACTGGAC 1598  
QY 867 GCGGGCGGAAACGTTGCGATCTGGAAGTAGGACAGGTCCGAGATC 912  
Db 1599 GCGGGCGGAAACGTTGCGATCTGGAAGTAGGACAGGTCCGAGATC 1644

RESULT 8  
AAQ05956  
ID AAQ05956 standard; DNA; 9185 BP.  
XX  
AC AAQ05956;  
XX

DT 25-MAR-2003 (revised)  
DT 23-JAN-1991 (first entry)  
XX  
DE Sense strand of the compiled Hepatitis C virus cDNA sequence.  
XX  
KW Hepatitis C virus (HCV); antiviral agent; ss.  
OS Hepatitis C virus.  
XX  
FH Key Location/Qualifiers  
FT misc\_RNA 1..1667  
FT /\*tag= b

```
FT CDS /note= "epitope within this region is claimed"
FT 320..9185
FT /*tag= a
FT 8978..9185
FT /*tag= c
FT /note= "encodes an epitope that is claimed"
FT
XX EP388232-A.
XX
XX 19-SEP-1990.
XX
XX 16-MAR-1990; 90EP-00302866.
XX
XX 17-MAR-1989; 89US-00325338.
XX
XX 20-APR-1989; 89US-00341334.
XX
XX 18-MAY-1989; 89US-00355002.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo QL, Kuo G;
XX
XX WPI; 1990-284418/38.
XX
XX P-PSDB; AAR08124.
XX
XX Hepatitis C virus DNA - used for producing probes, polypeptide(s),
XX antibodies and anti-sense polynucleotide(s) for diagnosis and therapy.
XX
XX Disclosure; Fig 17; 83pp; English.
XX
XX HCV cDNA libraries were constructed using pooled serum from a chimpanzee
XX with chronic HCV infection. A lambda gt10 library was screened with
XX probes derived from previously isolated clones. The ORF is derived from
XX the overlapping clones b114a, ag30a, CA205a, CA290a, CA216a, p14a,
XX CA167b, CA156e, CA84a, CA59a, K9-1, 26j, 131, 12f, 141, 11b, 7f, 8h, 33c,
XX 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g, 39c, 35f, 19g, 26g
XX 15e, b5a and 16jh. These clones extend the sequence of the HCV genome
XX reported in EP-318216. The upstream region from nucleotides -319 to +1348
XX (=1-1667 in this file) is covered by clones b114a, 18g, ag30a, CA205a
XX CA290a, CA216a, p14a, CA167b, CA156e, CA84a and CA59a; nucleotides 8659-
XX 8866 (=8978-9185 in this file) are covered by clones b5a and 16jh. See
XX also AAQ05955. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;
XX
XX Query Match 52.4%; Score 834.8; DB 2; Length 9185;
XX Best Local Similarity 96.4%; Pred. No. 3.6e-133;
XX Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
XX
XX 27 CTGTGCTGCTGCTGTGGAGCAGTCTTCTGTTGCGCCAGCGCTAGCGAAACCCACGT 86
XX 1420 CTGGGCGAAGGTCCTGTGTAGTGTCTGTCTATTTGCGCGCTGCGAGCGGAACCCACGT 1479
XX
XX 87 CACCGGGGAAGTGGCGGCCACACTGTGTCTGGATTGTAGCCTCTCGCACCAAGGCGC 146
XX 1480 CACCGGGGAAGTGGCGGCCACACTGTGTCTGGATTGTAGCCTCTCGCACCAAGGCGC 1539
XX
XX 147 CAAAGCAGAACGTCAGCTGATCAACCAACGCGAGTGGCAGCTCAATAGCAGCGCCT 206
XX 1540 CAAAGCAGAACGTCAGCTGATCAACCAACGCGAGTGGCAGCTCAATAGCAGCGCCT 1599
XX
XX 207 GAACGCAATGATAGCTCAACACCGGCTGTTGGCAGGCTTTCTATCACCACAAGTT 266
XX 1600 GAACGCAATGATAGCTCAACACCGGCTGTTGGCAGGCTTTCTATCACCACAAGTT 1659
XX
XX 267 CAACTCTTCAGGCTGCTCTGAGAGGCTAGCCAGTCCGACCCCTTACCAGATTGACCA 326
XX 1660 CAACTCTTCAGGCTGCTCTGAGAGGCTAGCCAGTCCGACCCCTTACCAGATTGACCA 1719
XX
XX 327 GGGCTGGGCGCTTATCAGTTATGCCAACGAGCGGCGCCGACGCGCCCTACTGCTG 386
XX 1720 GGGCTGGGCGCTTATCAGTTATGCCAACGAGCGGCGCCGACGCGCCCTACTGCTG 1779
XX
XX 387 GCACACTACCCCAACCTTGGCGGTATTTGGCCCGGAAGAGTGTGTGTGTCGGTATA 446
XX
Db 1780 GCACACTACCCCAACCTTGGCGGTATTTGGCCCGGAAGAGTGTGTGTGTCGGTATA 1839
Qy 447 TTGCTTCACTCCAGAGCCCGTGGTGGGAACACGACACAGGTGCGGCGCCACCTA 506
Db 1840 TTGCTTCACTCCAGAGCCCGTGGTGGGAACACGACAGGTGCGGCGCCACCTA 1899
Qy 507 CAGCTGGGGTGAAAATGATACGGACGCTTCTCGTCTTTAAACAATACAGGCCACCGCTGG 566
Db 1900 CAGCTGGGGTGAAAATGATACGGACGCTTCTCGTCTTTAAACAATACAGGCCACCGCTGG 1959
Qy 567 CAATTGGTTGCGTTGTATCTGATGAACTAACTGGAATTCACCAAGTGTGCGGAGCGCC 626
Db 1960 CAATTGGTTGCGTTGTATCTGATGAACTAACTGGAATTCACCAAGTGTGCGGAGCGCC 2019
Qy 627 TCCTTGTGTCACTCGAGGGGGGCAACAACACCTGCACTGCCCACTGATTGCTTCG 686
Db 2020 TCCTTGTGTCACTCGAGGGGGGCAACAACACCTGCACTGCCCACTGATTGCTTCG 2079
Qy 687 CAAGCATCCGACGCCACATACCTCTCGGTGCGGCTCCGCTCCCTGGATCACACCCAGGTG 746
Db 2080 CAAGCATCCGACGCCACATACCTCTCGGTGCGGCTCCGCTCCCTGGATCACACCCAGGTG 2139
Qy 747 CCTGTCGACTACCCGTATAGGCTTTGGCATATCTTTGTACCATCAACTACCATATT 806
Db 2140 CCTGTCGACTACCCGTATAGGCTTTGGCATATCTTTGTACCATCAACTACCATATT 2199
Qy 807 TAAATCAGGATGTACGTGGGAGGCTCGAACACAGGCTGGAAGCTGCTGCACTGGAC 866
Db 2200 TAAATCAGGATGTACGTGGGAGGCTCGAACACAGGCTGGAAGCTGCTGCACTGGAC 2259
Qy 867 GCGGGCGAAACGTTGCGATCTGGAAGATAGGACAGAGTCCGAGATC 912
Db 2260 GCGGGCGAAACGTTGCGATCTGGAAGATAGGACAGAGTCCGAGATC 2305
XX
XX RESULT 9
XX AAQ10566
XX ID AAQ10566 standard; DNA; 9185 BP.
XX XX
XX AC AAQ10566;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 29-APR-1991 (first entry)
XX XX
XX XX Hepatitis C virus strain 1 DNA.
XX DE
XX XX Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;
XX KW
XX XX viral infections; ss.
XX XX
XX OS Hepatitis C virus.
XX XX
XX PN EP414475-A.
XX XX
XX PD 27-FEB-1991.
XX XX
XX PF 21-AUG-1990; 90EP-00309120.
XX XX
XX PR 25-AUG-1989; 89US-00398667.
XX XX
XX XX (CHIR ) CHIRON CORP.
XX XX
XX XX Weiner AJ, Steimer KS;
XX XX
XX XX WPI; 1991-059670/09.
XX XX
XX XX Cell lines infected with hepatitis C virus - are used as source of
XX PT antigens for detection of HCV antibodies, for vaccines, and for screening
XX PT anti-viral agents.
XX XX
XX PS Disclosure; Fig 1; 24pp; English.
XX XX
XX CC This is a hepatitis C virus (HCV) composite cDNA sequence, deduced using
```

CC overlapping clones. a compsn. contg. the antigenic protein encoded by  
CC this sequence is useful for detecting anti-HCV anti- bodies (Abs) and for  
CC screening an agent which inhibits HCV replic- ation. A cell line infected  
CC with this virus can be used as a source of antigens. The antigen is  
CC useful for preparing vaccines for treating viral infections. See also  
CC AAQ10567. (updated on 25-MAR-2003 to correct PA field.)  
XX

SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Query Match		52.4%;	Score 834.8;	DB 2;	Length 9185;
Best Local Similarity		96.4%;	Pred. No. 3.6e-133;		
Matches 854;		Conservative 0;	Mismatches 32;	Indels 0;	Gaps 0;
QY	27	CTGTGTGCTGCTGCTGTGTGGAGCAGTCTTTCGTTTCGCCCGAGCGGTAGCGAAACCCACGT	86		
DB	1420	CTGGGCGAAGTCTCGTAGTGTCTGCTGCTATTTGCGGGCGTGCAGCGGAAACCCACGT	1479		
QY	87	CACCGGGGGAAGTGGCGGCACACAGTGTGTCTGATTTGTAGCTTCCTCGACACGAGCGC	146		
DB	1480	CACCGGGGGAAGTGGCGGCACACAGTGTGTCTGATTTGTAGCTTCCTCGACACGAGCGC	1539		
QY	147	CAAGCAGAACGTCAGCTGATCAACACCAACCGCAGTTGGCACCTCAATAGCACGGCCCT	206		
DB	1540	CAAGCAGAACGTCAGCTGATCAACACCAACCGCAGTTGGCACCTCAATAGCACGGCCCT	1599		
QY	207	GAATCGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCAAGTT	266		
DB	1600	GAATCGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATCACCAAGTT	1659		
QY	267	CAACTCTTCAGGCTGCTGAGAGGTAGCAGCTGCGACCCCTTACGATTTTGACCA	326		
DB	1660	CAACTCTTCAGGCTGCTGAGAGGTAGCAGCTGCGACCCCTTACGATTTTGACCA	1719		
QY	327	GGGCTGGGGCCCTATCAGTTATGCCAAGCGAGCGGCCCGACAGCGCCCTACTGCTG	386		
DB	1720	GGGCTGGGGCCCTATCAGTTATGCCAAGCGAGCGGCCCGACAGCGCCCTACTGCTG	1779		
QY	387	GCACTACCCCCCAAAACCTTCGGGTATTGTGCGCGGAGAGTGTGTGTCGGGTATA	446		
DB	1780	GCACTACCCCCCAAAACCTTCGGGTATTGTGCGCGGAGAGTGTGTGTCGGGTATA	1839		
QY	447	TGCTTCTACTCCAGCGCCCGTGGTGGGAAACGACGAGTCGGGCGGCCACCTTA	506		
DB	1840	TGCTTCTACTCCAGCGCCCGTGGTGGGAAACGACGAGTCGGGCGGCCACCTTA	1899		
QY	507	CAGCTGGGGTGAATGATACGAGCTCTTCGTCTTAAACATACAGCGCACCGCTGGG	566		
DB	1900	CAGCTGGGGTGAATGATACGAGCTCTTCGTCTTAAACATACAGCGCACCGCTGGG	1959		
QY	567	CAATTGGTTGCGTTGCTGATGAACTCAACTGGATTTCACCAAGTGTGGGAGCGCC	626		
DB	1960	CAATTGGTTGCGTTGCTGATGAACTCAACTGGATTTCACCAAGTGTGGGAGCGCC	2019		
QY	627	TCCTTGTGTATCGAGGGCGGGCAACACCTGTGATGCCCACTGATTTGCTCG	686		
DB	2020	TCCTTGTGTATCGAGGGCGGGCAACACCTGTGATGCCCACTGATTTGCTCG	2079		
QY	687	CAAGCATCCGAGCGCACATCTCTCGGTGGGCTCCGTCCTCGATCACACCCAGGTG	746		
DB	2080	CAAGCATCCGAGCGCACATCTCTCGGTGGGCTCCGTCCTCGATCACACCCAGGTG	2139		
QY	747	CCTGTGCTACCTACCGGTATAGGCTTTGGCATTTATCTTGTACCATCACTACACCATAT	806		
DB	2140	CCTGTGCTACCTACCGGTATAGGCTTTGGCATTTATCTTGTACCATCACTACACCATAT	2199		
QY	807	TAAATCAGGATGACGTGGAGGGGTGAAACAGAGGTGGAAGCTGCTGCACTGAC	866		
DB	2200	TAAATCAGGATGACGTGGAGGGGTGAAACAGAGGTGGAAGCTGCTGCACTGAC	2259		
QY	867	GCGGGGCAACGTTGCGATCTGGAAGATAGGACAGAGTCCGAGATC	912		
DB	2260	GCGGGGCAACGTTGCGATCTGGAAGATAGGACAGAGTCCGAGATC	2305		

RESULT 10

AXX00459  
ID AAX00459 standard; DNA; 9185 BP.

XX AC AAX00459;

XX DT 10-MAY-1999 (first entry)

XX DE Hepatitis C virus 1 ORF1.

XX HCV; non-A, non-B hepatitis; NANBH; diagnosis; vaccine; antigen;  
KW antibody; immunoassay; assay; ss.

XX OS Hepatitis C virus.

XX FH Key Location/Qualifiers  
FT CDS 320..9185  
FT /\*tag= a

XX PN US5856437-A..

XX PD 05-JAN-1999.

XX PF 03-NOV-1994; 94US-00334255.

XX PR 15-SEP-1989; 89US-00408045.

XX PR 21-DEC-1989; 89US-00456142.

XX PR 04-JAN-1991; 91US-00637380.

XX PR 02-AUG-1993; 93US-00101280.

XX PR 24-FEB-1994; 94US-00201066.

XX PA (NAHE-) NAT INST OF HEALTH JAPAN.

XX PA (CHIR) CHIRON CORP.

XX PI Weiner AJ, Miyamura T, Irvine BD, Kolberg JA, Han J, Saito I;

XX PI Cha T, Houghton M;

XX PI WPI; 1999-105191/09.

XX PT Antigenic polypeptides from J1 and J7 hepatitis C virus isolates - useful  
PT as immunoassay reagents, for raising antibodies and as vaccine  
PT components.

XX PS Disclosure; Fig 12A-C; 44pp; English.

XX This is the nucleotide sequence of ORF1 of hepatitis C virus prototype  
CC HCV-1. The invention relates to 2 new isolates, J1 and J7, of HCV. These  
CC new isolates comprise nucleotide (see AAX00400-04) and amino acid (see  
CC AAX0583-87) sequences which are distinct from HCV1. Thus J1 and J7  
CC provide new polynucleotides and polypeptides useful in diagnostics,  
CC recombinant protein production and vaccine development. Claimed  
CC polypeptides can be used: (i) as immunoassay reagents, or standards, to  
CC detect HCV antibodies, e.g. for diagnosing infection or screening donated  
CC blood; (ii) to generate specific antibodies (used for detecting the  
CC corresponding polypeptide, to screen for antiviral agents, for virus  
CC isolation and for passive immunisation); (iii) in protective or  
CC therapeutic vaccines, and (iv) for isolation of non-A, non-B viruses

XX SQ Sequence 9185 BP; 1849 A; 2793 C; 2605 G; 1938 T; 0 U; 0 Other;

Query Match 52.4%; Score 834.8; DB 2; Length 9185;  
Best Local Similarity 96.4%; Pred. No. 3.6e-133;  
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGTGCTGCTGCTGTGTGGAGCAGTCTTTCGTTTCGCCCGAGCGGTAGCGAAACCCACGT 86  
DB 1420 CTGGGCGAAGTCTCGTAGTGTCTGCTATTTGCGGGCGTGCAGCGGAAACCCACGT 1479

QY 87 CACCGGGGGAAGTGGCGGCACACAGTGTGTCTGATTTGTAGCTTCCTCGACACGAGCGC 146  
DB 1480 CACCGGGGGAAGTGGCGGCACACAGTGTGTCTGATTTGTAGCTTCCTCGACACGAGCGC 1539

Qy	147	CAAGCAGAACTCCAGCTGATCAACACCAACGCGAGTTGGACCTCAATAGCAGCGCCT	206
Db	1540	CAAGCAGAACTCCAGCTGATCAACACCAACGCGAGTTGGACCTCAATAGCAGCGCCT	1599
Qy	207	GAACTGCAATAGATAGCTCAACACCGCTGTTGGCAGGGCTTTTCTATACCAACAAGTT	266
Db	1600	GAACTGCAATAGATAGCTCAACACCGCTGTTGGCAGGGCTTTTCTATACCAACAAGTT	1659
Qy	267	CAACTCTTCAGGCTGCTCTGAGAGCTAGCCAGCTGCGACCCCTTACCGATTTTGACCA	326
Db	1660	CAACTCTTCAGGCTGCTCTGAGAGCTAGCCAGCTGCGACCCCTTACCGATTTTGACCA	1719
Qy	327	GGGCTGGGGCCCTATCAGTTATGCCAAGCGGGCCCGACAGCGCCCTACTGCTG	386
Db	1720	GGGCTGGGGCCCTATCAGTTATGCCAAGCGGGCCCGACAGCGCCCTACTGCTG	1779
Qy	387	GCACTACCCCCCAAACTTCGGGTATTGTCGCGGAAGAGTGTTGTTGTCGGGTATA	446
Db	1780	GCACTACCCCCCAAACTTCGGGTATTGTCGCGCGAAGAGTGTTGTTGTCGGGTATA	1839
Qy	447	TTGCTTCACTCCAGCCCGTGGTGGGAAACGACAGCTGCGACCCCTTACCGATTTTGACCA	506
Db	1840	TTGCTTCACTCCAGCCCGTGGTGGGAAACGACAGCTGCGACCCCTTACCGATTTTGACCA	1899
Qy	507	CAGCTGGGGTGAATGATAGCAGCTCTTGGTCAATACCAATACCGCCACCGCTGGG	566
Db	1900	CAGCTGGGGTGAATGATAGCAGCTCTTGGTCAATACCAATACCGCCACCGCTGGG	1959
Qy	567	CAATTTGGTTCGGTCTGCTGATCAACTCACTGATTTTCACTCAAGTGTGCGAGCGCC	626
Db	1960	CAATTTGGTTCGGTCTGCTGATCAACTCACTGATTTTCACTCAAGTGTGCGAGCGCC	2019
Qy	627	TCTTGTGTATCGGAGGGGGGCAACACACCTCGCATCTGCCCACTGATTTGTTCCG	686
Db	2020	TCTTGTGTATCGGAGGGGGGCAACACACCTCGCATCTGCCCACTGATTTGTTCCG	2079
Qy	687	CAAGCATCGGACCGCAATATCTCGTGGGGCTCCGGTCCCTGGATCAGACCCAGTG	746
Db	2080	CAAGCATCGGACCGCACATCTCTCGTGGGGCTCCGGTCCCTGGATCAGACCCAGTG	2139
Qy	747	CTGTGTCGACTACCCGTATAGGCTTTGGCATTTATCTTGTACCATCACTACCATATT	806
Db	2140	CTGTGTCGACTACCCGTATAGGCTTTGGCATTTATCTTGTACCATCACTACCATATT	2199
Qy	807	TAAATCAGGATGTAGTGGGAGGGTCGAACACAGGCTGGAAGCTGCACTGCACTGGAC	866
Db	2200	TAAATCAGGATGTAGTGGGAGGGTCGAACACAGGCTGGAAGCTGCACTGCACTGGAC	2259
Qy	867	GCGGGCGAAACGTTGCGATCTGGAAGATAGGACAGTCCGAGATC	912
Db	2260	GCGGGCGAAACGTTGCGATCTGGAAGATAGGACAGGTCGAGCTC	2305
RESULT 11			
ID	AAx26737	AAx26737 standard; DNA; 9185 BP.	
XX	AC	AAx26737;	
XX	DT	21-JUN-1999 (first entry)	
XX	DE	Nucleotide sequence of the ORF of HCV1.	
XX	KW	HCV; J7 isolate; J1 isolate; HCV1; immunoassay; asiatic strain;	
XX	KW	diagnosis; HCV infection; blood screening; immunisation; antiviral; ss.	
XX	OS	Hepatitis C virus.	
XX	PN	US871903-A.	
XX	PD	16-FEB-1999.	
XX	PF	08-MAY-1995; 95US-00436965.	

XX	15-SEP-1989;	89US-00408045.
PR	21-DEC-1989;	89US-00456142.
PR	04-JAN-1991;	91US-00637380.
PR	02-AUG-1993;	93US-00101280.
PR	24-FEB-1994;	94US-00201066.
PR	03-NOV-1994;	94US-00334255.
XX	(NAHE-) NAT INST OF HEALTH JAPAN.	
PA	(CHIR ) CHIRON CORP.	
XX	Saito I, Miyamura T;	
XX	WPI; 1999-166619/14.	
XX	Immunosays for Asiatic strains of hepatitis C virus - for diagnosis of	
PT	infection and screening blood supplies.	
PT	Disclosure; Fig 12A-C; 43pp; English.	
XX	The present sequence represents the sequence of ORF1 of hepatitis C virus	
CC	(HCV) isolate HCV1. The specification describes new isolates of HCV, J1	
CC	and J7. J1 and J7 isolates comprise sequences which are distinct from the	
CC	prototype HCV isolates, HCV1. The specific specification describes immunoassays	
CC	for HCV based on antigens from Asiatic strains not cross-reactive with	
CC	HCV-1. The assays are used for diagnosis of HCV infection and to screen	
CC	donated blood. The anti-HCV antibodies are also useful therapeutically	
CC	and prophylactically (passive immunisation); in screening for anti-viral	
CC	agents; for isolation, purification and identification of non-A, non-B	
CC	hepatitis virus (e.g. by affinity chromatography) and to raise anti-	
CC	idiotypic antibodies (useful for treatment or diagnosis) and to determine	
CC	immunogenic regions of the HCV antigens)	
XX	Sequence 9185 BP; 1849 A; 2792 C; 2606 G; 1938 T; 0 U; 0 Other;	
SQ	Query Match 52.4%; Score 834.8; DB 2; Length 9185;	
	Best Local Similarity 96.4%; Pred. No. 3.6e-133;	
	Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;	
Qy	27	CTGTGTGCTGCTGCTGTGGAGCAGCTCTCGTTTGGCCCGAGCGCTAGCGAAACCCACGT 86
Db	1420	CTGGGCGAAGTCTCTGGTGTGCTGCTATTTTGGCCGGCTGAGCGGAAACCCACGT 1479
Qy	87	CACCGGGGAAAGTGGCGGCCACACTGTGTCTGGATTGTGTAGCTCTCTCGACACGAGCGC 146
Db	1480	CACCGGGGAAAGTGGCGGCCACACTGTGTCTGGATTGTGTAGCTCTCTCGACACGAGCGC 1539
Qy	147	CAAGCAGAACGTCAGCTGATCAACACCAACGCGAGTTGGACCTCAATAGCAGCGCCT 206
Db	1540	CAAGCAGAACGTCAGCTGATCAACACCAACGCGAGTTGGACCTCAATAGCAGCGCCT 1599
Qy	207	GAACTGCAATAGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATACCAACAAGTT 266
Db	1600	GAACTGCAATAGATAGCTCAACACCGGCTGTTGGCAGGGCTTTTCTATACCAACAAGTT 1659
Qy	267	CAACTCTTCAGGCTGCTCTGAGAGCTAGCCAGCTGCGACCCCTTACCGATTTTGACCA 326
Db	1660	CAACTCTTCAGGCTGCTCTGAGAGCTAGCCAGCTGCGACCCCTTACCGATTTTGACCA 1719
Qy	327	GGGCTGGGGCCCTATCAGTTATGCCAAGCGGGCCCGACAGCGCCCTACTGCTG 386
Db	1720	GGGCTGGGGCCCTATCAGTTATGCCAAGCGGGCCCGACAGCGCCCTACTGCTG 1779
Qy	387	GCACTACCCCCCAAAACCTTGGGTATTGTCGCGGAAGAGTGTTGTTGTCGGGTATA 446
Db	1780	GCACTACCCCCCAAAACCTTGGGTATTGTCGCGCGAAGAGTGTTGTTGTCGGGTATA 1839
Qy	447	TTGCTTCACTCCAGCCCGTGGTGGGAAACGACAGCTGCGACCCCTTACCGATTTTGACCA 506
Db	1840	TTGCTTCACTCCAGCCCGTGGTGGGAAACGACAGCTGCGACCCCTTACCGATTTTGACCA 1899
Qy	507	CAGCTGGGGTGAATGATAGCAGCTCTTGGTCAATACCAATACCGCCACCGCTGGG 566

Db 1900 CAGCTGGGGTGAATATGATACGGACGCTCTTCCTCTTAAACAATACAGGCCACCGCTGGG 1959  
QY 567 CAATTGGTTTCGGTTGTACCTGATGAACCTCAACTGATTCACAAAGTGTGGGAGCGCC 626  
Db 1960 CAATTGGTTTCGGTTGTACCTGATGATCACTCAACTGATTCACAAAGTGTGGGAGCGCC 2019  
QY 627 TCCTTGTGTATCGAGGGGCGGGCAACACACCTGCACTGCCCCACTGATTTGCTTCGG 686  
Db 2020 TCCTTGTGTATCGAGGGGCGGGCAACACACCTGCACTGCCCCACTGATTTGCTTCGG 2079  
QY 687 CAAGCATCCGGACGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCAGACCCAGGTG 746  
Db 2080 CAAGCATCCGGACGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCAGACCCAGGTG 2139  
QY 747 CCTGGTCCAGTACCCGCTATAGGCTTTGGCATTTATCTGTACCATCAACTACACCATATT 806  
Db 2140 CCTGGTCCAGTACCCGCTATAGGCTTTGGCATTTATCTGTACCATCAACTACACCATATT 2199  
QY 807 TAAATCAGGATGTACGTGGGAGGGTGAACACAGGCTGGAGCTGCAACTGGAC 866  
Db 2200 TAAATCAGGATGTACGTGGGAGGGTGAACACAGGCTGGAGCTGCAACTGGAC 2259  
QY 867 GCGGGGCGAAGCTTCGGATCTGGAAGATAGGACAGGCTCCGAGATC 912  
Db 2260 GCGGGGCGAAGCTTCGGATCTGGAAGATAGGACAGGCTCCGAGCTC 2305

RESULT 12

ADPF6068

ID ADF66068 standard; DNA; 9185 BP.

XX AC

XX ADF66068;

XX 12-FEB-2004 (first entry)

XX Hepatitis C virus 1 ORF nucleotide sequence.

DE Hepatitis C virus; HCV; screening; identification; non A non B hepatitis;

KW NANBH; gene; de.

XX Hepatitis C virus.

OS Hepatitis C virus.

XX US959092-A.

PN 28-SEP-1999.

XX 08-MAY-1995; 95US-00436966.

XX 15-SEP-1989; 89US-00408045.

PR 21-DEC-1989; 89US-00456142.

PR 04-JAN-1991; 91US-00637380.

PR 02-AUG-1993; 93US-00101280.

PR 24-FEB-1994; 94US-00201066.

PR 03-NOV-1994; 94US-00334255.

XX (CHIR ) CHIRON CORP.

PA (NAHE-) NAT INST OF HEALTH JAPAN.

XX Saito I, Miyamura T;

XX WPI; 1999-570774/48.

XX Polynucleotide probes useful for screening and identifying carriers of

PT non A non B hepatitis.

XX Disclosure; Fig 12; 44pp; English.

CC The present invention describes polynucleotides sequences from Hepatitis

CC C virus (HCV) isolates J7 C/E, J1 E, J1 E/NS1, J1 NS3, and J1 NS5 or from

CC deposits BP-2593, BP2594, BP-2595, BP-2637 and BP-2638. The

CC polynucleotide sequences are useful as probes for screening and

CC identifying carriers of non A non B hepatitis (NANBH). The present

CC sequence is used in the exemplification of the present invention.

XX SQ Sequence 9185 BP; 1849 A; 2792 C; 2606 G; 1938 T; 0 U; 0 Other;  
Query Match 52.4%; Score 834.8; DB 2; Length 9185;  
Best Local Similarity 96.4%; Pred. No. 3.6e-133;  
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
QY 27 CTGTGTCTGTCTGTCTGTGGAGAGTCTTCGTTTCGCCCGCAGCGCTAGCGAAACCCACCT 86  
Db 1420 CTGGGCGAAGGTCCTGGTAGTGTCTCTATTTCCCGCGCTCGAGCGGAAACCCACCT 1479  
QY 87 CACCGGGGAAAGTGCCTGCCCAACATCTGTCTGTGATTTGTATGCTCTCTCGACCGAGCGC 146  
Db 1480 CACCGGGGAAAGTGCCTGCCCAACATCTGTCTGTGATTTGTATGCTCTCTCGACCGAGCGC 1539  
QY 147 CAACGAGACGTCAGCTGATCAACACCAACCGGAGTTGGACCTCAATAGCAGCGCCCT 206  
Db 1540 CAACGAGACGTCAGCTGATCAACACCAACCGGAGTTGGACCTCAATAGCAGCGCCCT 1599  
QY 207 GAACCTGCAATAGTACCTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACAAGTT 266  
Db 1600 GAACCTGCAATAGTACCTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACAAGTT 1659  
QY 267 CAACCTTTTCAAGCTGTCTGTAGAGCTAGCCAGCTGCCAGCCCTTACCGATTTTGAACA 326  
Db 1660 CAACCTTTTCAAGCTGTCTGTAGAGCTAGCCAGCTGCCAGCCCTTACCGATTTTGAACA 1719  
QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACAGCGCCCTACTGCTG 386  
Db 1720 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGACAGCGCCCTACTGCTG 1779  
QY 387 GCACCTACCCCCCAAAACCTTCGGGTATTGTGCCCGAAGAGTGTGTGGTCCGGTATA 446  
Db 1780 GCACCTACCCCCCAAAACCTTCGGGTATTGTGCCCGAAGAGTGTGTGGTCCGGTATA 1839  
QY 447 TTGCTTCACTCCCAAGCCCGTGGTGGGAAACGACAGAGTCCGGCGCGCCACCTA 506  
Db 1840 TTGCTTCACTCCCAAGCCCGTGGTGGGAAACGACAGAGTCCGGCGCGCCACCTA 1899  
QY 507 CAGCTGGGGTGAAATGATACGGACGCTTCGTCTTAAACATACCGAGCCACCGCTGGG 566  
Db 1900 CAGCTGGGGTGAAATGATACGGACGCTTCGTCTTAAACATACCGAGCCACCGCTGGG 1959  
QY 567 CAATTGGTTTCGGTTGTACCTGATGAACCTCAACTGGATTACCAAAAGTGTGGGAGCGCC 626  
Db 1960 CAATTGGTTTCGGTTGTACCTGATGAACCTCAACTGGATTACCAAAAGTGTGGGAGCGCC 2019  
QY 627 TCCTTGTGTATCGAGGGGCGGGCAACACACCTGCACTGCCCCACTGATTTGCTTCG 686  
Db 2020 TCCTTGTGTATCGAGGGGCGGGCAACACACCTGCACTGCCCCACTGATTTGCTTCG 2079  
QY 687 CAAGCATCCGGACGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCAGACCCAGGTG 746  
Db 2080 CAAGCATCCGGACGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCAGACCCAGGTG 2139  
QY 747 CTGTGTGACTACCCGCTATAGGCTTTGGCATTTATCTCTGTATCACTCAACTACCATATT 806  
Db 2140 CTGTGTGACTACCCGCTATAGGCTTTGGCATTTATCTCTGTATCACTCAACTACCATATT 2199  
QY 807 TAAATCAGGATGTACGTGGGAGGGTGAACACAGGCTGGAAGCTGCTTCAACTGGAC 866  
Db 2200 TAAATCAGGATGTACGTGGGAGGGTGAACACAGGCTGGAAGCTGCTTCAACTGGAC 2259  
QY 867 GCGGGGCGAAGCTTCGGATCTGGAAGATAGGACAGGCTCCGAGATC 912  
Db 2260 GCGGGGCGAAGCTTCGGATCTGGAAGATAGGACAGGCTCCGAGCTC 2305

RESULT 13

AAA75297

ID AAA75297 standard; cDNA; 9185 BP.

XX

AC AAA75297;





XX (CHIR ) CHIRON CORP.  
 XX Houghton M, Choo Q, Kuo G;  
 XX WPI: 1996-117956/13.  
 XX P-PSDB; AAR90931.  
 XX  
 XX Combinations of synthetic Hepatitis C Virus antigens - provide more  
 XX effective diagnosis of Non-A, Non-B Hepatitis.  
 XX  
 XX Disclosure; Fig 1(A-Y); 53pp; English.  
 XX  
 XX The combination comprises an HCV antigen from the C domain (pref. C22 -  
 CC AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -  
 CC AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5  
 CC (AAR90934) domain. The antigens may in the form of a fusion protein, a  
 CC simple physical mixture, or the individual antigens commonly bound to a  
 CC solid matrix. They are pref. prep'd. by recombinant DNA techniques  
 CC (primers are given in AAT12711-T12716), but can be synthesised or  
 CC isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003  
 CC to correct PF field.)  
 XX  
 XX Sequence 9401 BP; 1893 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;  
 SQ  
 Query Match 52.4%; Score 834.8; DB 2; Length 9401;  
 Best Local Similarity 96.4%; Pred. No. 3.6e-133;  
 Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
 QY 27 CTGTGCTGCTGCTGTGGAGCAGTCTGCTTTCGCCCGGCTAGGCAACCCAGCT 86  
 DB 1442 CTGGCGAAGGTCCTGTGTCTGCTATTTGCGCGGCTCGACGGGAAACCCACGT 1501  
 QY 87 CACCGGGGAAGTGGCGGCACTGTGCTGATTTGTAGCTCTCGACACAGGGCG 146  
 DB 1502 CACCGGGGAAGTGGCGGCACTGTGCTGATTTGTAGCTCTCGACACAGGGCG 1561  
 QY 147 CAAGCAGAACGTCAGCTGATCAACACCAACGGCAGTTGGCACCTCAATAGCAGCGCCCT 206  
 DB 1562 CAAGCAGAACGTCAGCTGATCAACACCAACGGCAGTTGGCACCTCAATAGCAGCGCCCT 1621  
 QY 207 GAACCTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACAAGTT 266  
 DB 1622 GAACCTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACAAGTT 1681  
 QY 267 CAACCTCTTCAGGCTGCTGAGAGCTAGCAGCTGCGACCCCTTACCGATTTTGACCA 326  
 DB 1682 CAACCTCTTCAGGCTGCTGAGAGCTAGCAGCTGCGACCCCTTACCGATTTTGACCA 1741  
 QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGCGAGCGGCCCGACAGCGCCCTACTGCTG 386  
 DB 1742 GGGCTGGGGCCCTATCAGTTATGCCAAGCGAGCGGCCCGACAGCGCCCTACTGCTG 1801  
 QY 387 GCATACCCCCCAAAACCTTCGGGTATGTTGCCCCGAGAGTGTGTGTGTCGGGTATA 446  
 DB 1802 GCATACCCCCCAAAACCTTCGGGTATGTTGCCCCGAGAGTGTGTGTGTCGGGTATA 1861  
 QY 447 TTGCTTCACTCCAGCCCGTGGTGGGAACGACGACAGGTGGGGCGGCCACCTA 506  
 DB 1862 TTGCTTCACTCCAGCCCGTGGTGGGAACGACGACAGGTGGGGCGGCCACCTA 1921  
 QY 507 CAGCTGGGGTGAATATGATACGAGCTCTTCTGCTTAAACAATACAGGCCACCGCTGGG 566  
 DB 1922 CAGCTGGGGTGAATATGATACGAGCTCTTCTGCTTAAACAATACAGGCCACCGCTGGG 1981  
 QY 567 CAATTGGTTCGGTTGCTGATGAACTCAACTGGATTCAACAAAGTGTGCGAGCGCC 626  
 DB 1982 CAATTGGTTCGGTTGCTGATGAACTCAACTGGATTCAACAAAGTGTGCGAGCGCC 2041  
 QY 627 TCCTTGTGTCTCGAGGGGGGCAACACACCTGCACTGCCCACTGATTTGCTCGG 686  
 DB 2042 TCCTTGTGTCTCGAGGGGGGCAACACACCTGCACTGCCCACTGATTTGCTCGG 2101

QY 687 CAAGCATCCGAGCGCCACATACTCTCGTGGCGCTCCGGTCCCTGGATCACACCCAGGTG 746  
 DB 2102 CAAGCATCCGAGCGCCACATACTCTCGTGGCGCTCCGGTCCCTGGATCACACCCAGGTG 2161  
 QY 747 CTGTGTCGACTACCCGTATAGGCTTTGGCATTATCTTGTATCCATCAACTACACCATATT 806  
 DB 2162 CTGTGTCGACTACCCGTATAGGCTTTGGCATTATCTTGTATCCATCAACTACACCATATT 2221  
 QY 807 TAAATATCAGATGTAGCTGGGAGGGTCCGAACACAGGCTGGAGCTCCCTGCAACTGGAC 866  
 DB 2222 TAAATATCAGATGTAGCTGGGAGGGTCCGAACACAGGCTGGAGCTCCCTGCAACTGGAC 2281  
 QY 867 GCGGGCGGAACGTTGCGATCTCGAAGATAGGACAGGTCGCGAGATC 912  
 DB 2282 GCGGGCGGAACGTTGCGATCTCGAAGATAGGACAGGTCGCGAGATC 2327  
 RESULT 15  
 AAT99981  
 ID AAT99981 standard; DNA; 9401 BP.  
 XX  
 AC AAT99981;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 16-MAR-1998 (first entry)  
 XX  
 DE HCV polyprotein coding sequence.  
 XX  
 KW PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;  
 KW C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;  
 KW NS4; ds.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 342..9377  
 FT /\*tag= a  
 XX  
 PN US5683864-A.  
 XX  
 PD 04-NOV-1997.  
 XX  
 PF 07-JUL-1992; 92US-00910760.  
 XX  
 PR 18-NOV-1987; 87US-00122714.  
 PR 30-DEC-1987; 87US-00139886.  
 PR 26-FEB-1988; 88US-00161072.  
 PR 06-MAY-1988; 88US-00191263.  
 PR 26-OCT-1988; 88US-00263584.  
 PR 14-NOV-1988; 88US-00271450.  
 PR 17-MAR-1989; 89US-00325338.  
 PR 20-APR-1989; 89US-00341334.  
 PR 21-APR-1989; 89US-00353896.  
 PR 18-MAY-1989; 89US-00355002.  
 PR 04-APR-1990; 90US-00504352.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Kuo G, Houghton M, Choo Q;  
 XX  
 DR WPI; 1997-548976/50.  
 DR P-PSDB; AAW34480.  
 XX  
 PT Combination of three hepatitis C virus antigens - used for detection of  
 PT specific antibodies to diagnose infection.  
 XX  
 PS Disclosure; Col 25-46; 57pp; English.  
 XX  
 CC This sequence represents the Hepatitis C virus polyprotein coding  
 CC sequence. Fragments of this sequence can be amplified and used in the  
 CC combination of HCV antigens of the invention. The HCV antigen combination  
 CC comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)  
 CC 1-120 of the HCV polyprotein), or its immunologically reactive fragment



CC containing at least 8 aa. It also comprises two additional antigens from  
CC two different polyprotein domains, including at least 8 aa from the NS3,  
CC NS4, S or NS5 domains of the polyprotein, corresponding, respectively, to  
CC aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein.  
CC Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa  
CC regions of the HCV polyprotein. These antigen combinations are used  
CC diagnostically to detect anti-HCV antibodies, using any standard  
CC immunoassay format. These antigen combinations have a broader range of  
CC reactivity with antibodies than any antigen individually. (Updated on 25-  
CC MAR-2003 to correct PR field.)  
XX  
SQ Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

Query Match 52.4%; Score 834.8; DB 2; Length 9401;  
Best Local Similarity 96.4%; Pred. No. 3.6e-133;  
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGTGCTGCTGCTGTGGAGCAGTCTCGTTTCGCCAGCGCTAGCGAAACCCACGT 86  
DB 1442 CTGGGCGAAGGTCCTGGTAGTGCTGCTGCTATTTGCCGGGTGAGCGGAAACCCACGT 1501

QY 87 CACCGGGGGAAGTCCGGCCACACACTGTGTCTGGATTGTAGCTTCCTCGCACAGGCGC 146  
DB 1502 CACCGGGGGAAGTCCGGCCACACACTGTGTCTGGATTGTAGCTTCCTCGCACAGGCGC 1561

QY 147 CAAGCAGAAAGTCCAGCTGATCAACCAACCGGCAAGTTGGCACCTCAATAGCACGGCCCT 206  
DB 1562 CAAGCAGAAAGTCCAGCTGATCAACCAACCGGCAAGTTGGCACCTCAATAGCACGGCCCT 1621

QY 207 GAACCTGCAATGATAGCCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCACAAGTT 266  
DB 1622 GAACCTGCAATGATAGCCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCACAAGTT 1681

QY 267 CAACCTCTTCAGGCTGTCTGAGAGCTAGCCAGCTGCCAGCCCTTACCGATTTTGACCA 326  
DB 1682 CAACCTCTTCAGGCTGTCTGAGAGCTAGCCAGCTGCCAGCCCTTACCGATTTTGACCA 1741

QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAACCGGAGCGCCCGGACAGCGCCCTACTGCTG 386  
DB 1742 GGGCTGGGGCCCTATCAGTTATGCCAAACCGGAGCGCCCGGACAGCGCCCTACTGCTG 1801

QY 387 GCACCTACCCCCCAAAACCTTGCGGTATTTGTCGCCGAAGAGTGTGTGGTCCGGTATA 446  
DB 1802 GCACCTACCCCCCAAAACCTTGCGGTATTTGTCGCCGAAGAGTGTGTGGTCCGGTATA 1861

QY 447 TTGCTTTCACTCCAGCCCGTGGTGGGAAACGACAGGTGGGGCGGCCCAACCTTA 506  
DB 1862 TTGCTTTCACTCCAGCCCGTGGTGGGAAACGACAGGTGGGGCGGCCCAACCTTA 1921

QY 507 CAGCTGGGGTGAATGATACGAGCTCTTCGTCTTAAACATACGAGCCACCGTGGG 566  
DB 1922 CAGCTGGGGTGAATGATACGAGCTCTTCGTCTTAAACATACGAGCCACCGTGGG 1981

QY 567 CAATTGGTTTCGGTTGTAACCTGGATGAACCTCAACTGGATTACCAAGTGTGGGAGCGCC 626  
DB 1982 CAATTGGTTTCGGTTGTAACCTGGATGAACCTCAACTGGATTACCAAGTGTGGGAGCGCC 2041

QY 627 TCCTTTGTGTATCGGAGGGCGGGCAACACACCTTCGACTGCCCACTGATTTGCTTCG 686  
DB 2042 TCCTTTGTGTATCGGAGGGCGGGCAACACACCTTCGACTGCCCACTGATTTGCTTCG 2101

QY 687 CAAGCATCCGGAGCGCCACATATCTCTCGGTGGGCTCCGCTCCCTGGATCACACCCAGGTG 746  
DB 2102 CAAGCATCCGGAGCGCCACATATCTCTCGGTGGGCTCCGCTCCCTGGATCACACCCAGGTG 2161

QY 747 CCTGGTCCGACTACCCGTATAGGCTTTGGCATTTATCTGTACCATCAACTACACCATATT 806  
DB 2162 CCTGGTCCGACTACCCGTATAGGCTTTGGCATTTATCTGTACCATCAACTACACCATATT 2221

QY 807 TAAATCAGGATGATAGCTGGGAGGGGTGGAACAACAGGCTGGAAGCTGCTGCAACTGGAC 866  
DB 2222 TAAATCAGGATGATAGCTGGGAGGGGTGGAACAACAGGCTGGAAGCTGCTGCAACTGGAC 2281

QY 867 GCGGGGCGAACGTTTGGATCTGGAAGATAGGGACAGGTCCGAGATC 912  
DB 2282 GCGGGGCGAACGTTTGGATCTGGAAGACAGGGACAGGTCCGAGCTC 2327

Search completed: April 14, 2005, 23:14:07  
Job time : 900 secs



Db 2292 TCCGACCCCTTACCGATTTTGACACAGGCTGGGCCCTATCAGTTATGCCAACCGAAGC 2351  
QY 361 GGGCCCGACAGCGGCCCTACTGCTGGCACTACCCCCCAAAACCTTGGGTATTTGGCC 420  
Db 2352 GGGCCCGACAGCGGCCCTACTGCTGGCACTACCCCCCAAAACCTTGGGTATTTGGCC 2411  
QY 421 GCGAAGAGTGTGTGGTCCGGTATATTGCTTCACTCCAGCCCGTGTGTGGGAAG 480  
Db 2412 GCGAAGAGTGTGTGGTCCGGTATATTGCTTCACTCCAGCCCGTGTGTGGGAAG 2471  
QY 481 ACCGACAGTCCGGCGGCCCACTACAGCTCGGGTGAAAAATGATACGACGCTTTCGTC 540  
Db 2472 ACCGACAGTCCGGCGGCCCACTACAGCTCGGGTGAAAAATGATACGACGCTTTCGTC 2531  
QY 541 CTTAAACAATACAGGCCACCGCTGGGCAATGGTTTCGGTTGTACCTGGATGAACCTCAACT 600  
Db 2532 CTTAAACAATACAGGCCACCGCTGGGCAATGGTTTCGGTTGTACCTGGATGAACCTCAACT 2591  
QY 601 GGATTACCAAAAGTGTGGAGCGCTCCTTGTGTATCGGAGGGCGGGCAACACACC 660  
Db 2592 GGAATTACCAAAAGTGTGGAGCGCTCCTTGTGTATCGGAGGGCGGGCAACACACC 2651  
QY 661 CTGCACTGCCCACTGATTGCTTCGCAAGCATCCGACGCCACATACCTCGGTGCGGC 720  
Db 2652 CTGCACTGCCCACTGATTGCTTCGCAAGCATCCGACGCCACATACCTCGGTGCGGC 2711  
QY 721 TCCGGTCCCTGGATCACACCCAGGTGCTGGTCCGACTACCCGTATAGGCTTTGGCAATTAT 780  
Db 2712 TCCGGTCCCTGGATCACACCCAGGTGCTGGTCCGACTACCCGTATAGGCTTTGGCAATTAT 2771  
QY 781 CCTTGTACCATCACTACACCATATTTAAATCAGGATGTAGTGGAGGGGTGCAACAC 840  
Db 2772 CCTTGTACCATCACTACACCATATTTAAATCAGGATGTAGTGGAGGGGTGCAACAC 2831  
QY 841 AGSCTGGAAGTCCCTGCAACTGGACGGGGCGAAAGTTGCGATCTGGAAGATAGGAC 900  
Db 2832 AGSCTGGAAGTCCCTGCAACTGGACGGGGCGAAAGTTGCGATCTGGAAGATAGGAC 2891  
QY 901 AGSCTCGAGATCGATATGAGAAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTGA 960  
Db 2892 AGSCTCGAGATCGATATGAGAAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTGA 2951  
QY 961 CAGCGGGGGTTTTCTTGTGTGACAGAATCTTCACAAATACCGGAGGTCTAGACTCGTG 1020  
Db 2952 CAGCGGGGGTTTTCTTGTGTGACAGAATCTTCACAAATACCGGAGGTCTAGACTCGTG 3011  
QY 1021 TGGACTTCTCAATTTTCTAGGGGATCTCCGCTGTGCTTGGCCAAATTCGCAAGTCC 1080  
Db 3012 TGGACTTCTCAATTTTCTAGGGGATCTCCGCTGTGCTTGGCCAAATTCGCAAGTCC 3071  
QY 1081 CCAACCTCCAATCACTCACCAACCTCCTGTCTCAATTTGCTCGTGTATCGCTGGATG 1140  
Db 3072 CCAACCTCCAATCACTCACCAACCTCCTGTCTCAATTTGCTCGTGTATCGCTGGATG 3131  
QY 1141 TGTCTCGGGGTTTTATCATATCTCTTCATCCTGCTGCTATGCTCATCTCTTATTTG 1200  
Db 3132 TGTCTCGGGGTTTTATCATATCTCTTCATCCTGCTGCTATGCTCATCTCTTATTTG 3191  
QY 1201 GTTCTTCTGATTATCAAGGTATGTTGCCGTTTGTCTCTAATTCAGGATCAACAACA 1260  
Db 3192 GTTCTTCTGATTATCAAGGTATGTTGCCGTTTGTCTCTAATTCAGGATCAACAACA 3251  
QY 1261 ACCAGTACGGGACCATGCAAAACCTGCAAGCTCTGCTCAAGGCAACTCTATGTTTCCC 1320  
Db 3252 ACCAGTACGGGACCATGCAAAACCTGCAAGCTCTGCTCAAGGCAACTCTATGTTTCCC 3311  
QY 1321 TCATGTTGCTGTAACAAAACCTAGGATGAAATTCACCTGTATTCGATCCCATCGTCC 1380  
Db 3312 TCATGTTGCTGTAACAAAACCTAGGATGAAATTCACCTGTATTCGATCCCATCGTCC 3371  
QY 1381 TGGGCTTTTCGAAAATACCTATGGGAGTGGGCTCAGTCCGTTTCTCTTGGCTCAGTTTA 1440

Db 3372 TGGGCTTTTCGAAAATACCTATGGAGTGGGCTCAGTCCGTTTCTCTTGGCTCAGTTTA 3431  
QY 1441 CTAGTGCCATTTGTTTCACTGGTTCTAGGGCTTTCCCCCACTGTTTGGCTTTCAAGTATA 1500  
Db 3432 CTAGTGCCATTTGTTTCACTGGTTCTAGGGCTTTCCCCCACTGTTTGGCTTTCAAGTATA 3491  
QY 1501 TGGATGATGTGTATTTGGGGCCAAAGTCTGTACACATCGTGAGTCCCTTTTATACCGCTG 1560  
Db 3492 TGGATGATGTGTATTTGGGGCCAAAGTCTGTACACATCGTGAGTCCCTTTTATACCGCTG 3551  
QY 1561 TTACCAATTTCTTTTGTCTCTGGGTATACATT 1593  
Db 3552 TTACCAATTTCTTTTGTCTCTGGGTATACATT 3584

RESULT 2  
US-09-693-596-3  
; Sequence 3, Application US/09693596  
; Patent No. 6521423  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-lim  
; APPLICANT: Abriagnani, Sergio  
; APPLICANT: Chien, David  
; APPLICANT: Selby, Mark  
; APPLICANT: Glazer, Edward  
; TITLE OF INVENTION: Intracellular Production of Hepatitis C E1 and E2  
; FILE OF INVENTION: Truncated Polypeptides  
; FILE REFERENCE: 1378.002  
; CURRENT APPLICATION NUMBER: US/09/693,596  
; CURRENT FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 09/073,406  
; PRIOR FILING DATE: 1998-05-06  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1989  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1989)  
US-09-693-596-3

Query Match 52.5%; Score 836.4; DB 4; Length 1989;  
Best Local Similarity 96.5%; Pred. No. 2.5e-259;  
Matches 855; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 27 CTGTGTGCTGCTGTGTGGAGCAGTCTTCTGCTTTCGCCCGCGCTAGCGAAACCCACGT 86  
Db 12 CTGGGGGAAGTCTCTGGTAGTCTGCTGCTATTTGCCGCGTGGACGGGAAACCAAGT 71  
QY 87 CACCGGGGAAGTGC CGGCCACACCTGTGTCTGGAATTTGTAGCTCTCGCACCGGCGC 146  
Db 72 CACCGGGGAAGTGC CGGCCACACCTGTGTCTGGAATTTGTAGCTCTCGCACCGGCGC 131  
QY 147 CAGCAGAAAGTCCAGCTGATCAACCAACCGGAGTGGCAGCTCAATAGCAGCGCCCT 206  
Db 132 CAGCAGAAAGTCCAGCTGATCAACCAACCGGAGTGGCAGCTCAATAGCAGCGCCCT 191  
QY 207 GAACTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACAAGTT 266  
Db 192 GAACTGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTTCTATCACAAGTT 251  
QY 267 CAACTCTTCAAGCTGTCTGAGAGGCTAGCCAGTCCGACCCCTTACCGATTTGACCA 326  
Db 252 CAACTCTTCAAGCTGTCTGAGAGGCTAGCCAGTCCGACCCCTTACCGATTTGACCA 311  
QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGGACGAGCGCCCTACTGCTG 386  
Db 312 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCCGGACGAGCGCCCTACTGCTG 371  
QY 387 GCACTACCCCGGCAAAACCTTGGGTATTTGTGCGCGGAGAGTGTGTGTCGGGTATA 446



QY 747 CTTGTCGACTACCCGTATAGGCTTTGGCATTATCTTGTACCATCAACTACACCATATT 806  
 DB 805 CTTGTCGACTACCCGTATAGGCTTTGGCATTATCTTGTACCATCAACTACACCATATT 864  
 QY 807 TAAATCAGGATGTACGTGGAGGGTCCGAAACACAGGCTGGAAGCTGCCTGCAACTGGAC 866  
 DB 865 TAAATCAGGATGTACGTGGAGGGTCCGAAACACAGGCTGGAAGCTGCCTGCAACTGGAC 924  
 QY 867 GCGGGCGAAGCTTGGATCTCGAAGATAGGACAGGTCGAGATC 912  
 DB 925 GCGGGCGAAGCTTGGATCTCGAAGATAGGACAGGTCGAGATC 970

RESULT 4  
 US-08-325-630-12  
 ; Sequence 12, Application US/08325630  
 ; Patent No. 5750331  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MIYAMURA, TATSUO  
 ; APPLICANT: SAITO, IZUMU  
 ; APPLICANT: HARADA, SHIZUKO  
 ; APPLICANT: HONDA, YOSHIKAZU  
 ; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS C  
 ; NUMBER OF SEQUENCES: 28  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 ; CITY: Arlington  
 ; STATE: Virginia  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22202

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/325,630  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/956,993  
 FILING DATE: 06-OCT-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Oblon, No. 5750331man F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 4667-001-0  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 413-3000  
 TELEFAX: (703) 413-2220  
 TELEX: 24855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1207 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Hepatitis C virus  
 IMMEDIATE SOURCE:  
 CLONE: HCV1  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 2..1207  
 US-08-325-630-12

Query Match 52.4%; Score 834.8; DB 1; Length 1207;  
 Best Local Similarity 96.4%; Pred. No. 6e-259;  
 Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTCGTCTGCTGTGTGGAGCAGTCTTGGTTTCGCCCGCAGCGCTAGCGAAACCCACCT 86  
 DB 85 CTGGCGGAAGGTCCTGGTAGTCTGCTGCTATTTGCCGCGCTCGACGCGGAAACCCACCT 144  
 QY 87 CACCGGGGAAGTGGCGGCCACACCTGTGTCTGGATTGTTAGCTCTCTCGCACACCGCC 146  
 DB 145 CACCGGGGAAGTGGCGGCCACACCTGTGTCTGGATTGTTAGCTCTCTCGCACACCGCC 204  
 QY 147 CAAGCAGAACCTTCAGCTGATCAACCAACGCGAGTTGGCACCTCAATAGACGCGCCCT 206  
 DB 205 CAAGCAGAACCTTCAGCTGATCAACCAACGCGAGTTGGCACCTCAATAGACGCGCCCT 264  
 QY 207 GAACTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTCTATCACCACAAGTT 266  
 DB 265 GAACTGCAATGATAGCTCAACACCGGCTGTTGGCAGGGCTTTCTATCACCACAAGTT 324  
 QY 267 CAACTCTTCAGGCTGCTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTGTTGACCA 326  
 DB 325 CAACTCTTCAGGCTGCTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTGTTGACCA 384  
 QY 327 GGGCTGGGGCCCTATCAGTTATGCGAAGCGAGCGGCCCGACAGCGCCCTCTACTGTG 386  
 DB 385 GGGCTGGGGCCCTATCAGTTATGCGAAGCGAGCGGCCCGACAGCGCCCTCTACTGTG 444  
 QY 387 GCACCTACCCCGCAAAACCTTGGGTATTGTCGCCCGGAAGAGTGTGTGTCGCGTATA 446  
 DB 445 GCACCTACCCCGCAAAACCTTGGGTATTGTCGCCCGGAAGAGTGTGTGTCGCGTATA 504  
 QY 447 TTGCTTCACTCCCGAGCCCGCTGGTGGGAAACGACCGACAGCTCGGGCGCCCGCACCTTA 506  
 DB 505 TTGCTTCACTCCCGAGCCCGCTGGTGGGAAACGACCGACAGCTCGGGCGCCCGCACCTTA 564  
 QY 507 CAGCTGGGGTGAATATGATACGAGCTCTTCTCTTAACTACAGCGCAACCGCTGGG 566  
 DB 565 CAGCTGGGGTGAATATGATACGAGCTCTTCTCTTAACTACAGCGCAACCGCTGGG 624  
 QY 567 CAATTGGTTCGGTCTGCTGATGAACTCAACTGATTCACCAAGTGTGCGGAGCGCC 626  
 DB 625 CAATTGGTTCGGTCTGCTGATGAACTCAACTGATTCACCAAGTGTGCGGAGCGCC 684  
 QY 627 TCCTTGTGTCTCGGAGGGCGGCAACCAACCCCTGCACTGCCCGCCTGATTGCTTCG 686  
 DB 685 TCCTTGTGTCTCGGAGGGCGGCAACCAACCCCTGCACTGCCCGCCTGATTGCTTCG 744  
 QY 687 CAAGCATCGGAGCGCCACATCTCTGGTGGGCTCCGCTCCGTCCGTGATCACACCGAGTG 746  
 DB 745 CAAGCATCGGAGCGCCACATCTCTGGTGGGCTCCGCTCCGTGATCACACCGAGTG 804  
 QY 747 CCGTGTGCACTACCCGTATAGGCTTTGGCATTATCTTGTACCATCAACTACACCATATT 806  
 DB 805 CCGTGTGCACTACCCGTATAGGCTTTGGCATTATCTTGTACCATCAACTACACCATATT 864  
 QY 807 TAAATCAGGATGTACGTGGAGGGTCCGAAACACAGGCTGGAAGCTGCTGCAACTGGAC 866  
 DB 865 TAAATCAGGATGTACGTGGAGGGTCCGAAACACAGGCTGGAAGCTGCTGCAACTGGAC 924  
 QY 867 GCGGGCGAAGCTTGGATCTCGAAGATAGGACAGGTCGAGATC 912  
 DB 925 GCGGGCGAAGCTTGGATCTCGAAGATAGGACAGGTCGAGATC 970

RESULT 5  
 US-08-444-818-88  
 ; Sequence 88, Application US/08444818  
 ; Patent No. 6150087  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chien, David Y.  
 ; APPLICANT: Rutter, William J.  
 ; TITLE OF INVENTION: NANV Diagnostics and Vaccines  
 ; NUMBER OF SEQUENCES: 777  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Chiron Corporation  
 ; STREET: 4560 Horton Street



```
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 27 CTGTGTGCTGCTGTGTGAGAGTCTTCTGTTTCGCCCCAGCGTAGCGAAACCCACGT 86
Db 1420 CTGGCGAAAGTCTCTGTAGTGTCTGTCTATTTGCGCGCTCGAGCGGAAACCCACGT 1479
QY 87 CACCGGGGAAGTGGCGGCCACACTGTGTCTGGATTGTTAGCTTCTCGCACAGGGCG 146
Db 1480 CACCGGGGAAGTGGCGGCCACACTGTGTCTGGATTGTTAGCTTCTCGCACAGGGCG 1539
QY 147 CAAGCAGAACGCTCCAGCTGATCAACACCAACGCGAGTTGGCAGCTCAATAGCACGGCCCT 206
Db 1540 CAAGCAGAACGCTCCAGCTGATCAACACCAACGCGAGTTGGCAGCTCAATAGCACGGCCCT 1599
QY 207 GAATCTCAATATAGCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCAAGTT 266
Db 1600 GAATCTCAATATAGCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCAAGTT 1659
QY 267 CAATCTTTCAGGCTGTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTGTGACCA 326
Db 1660 CAATCTTTCAGGCTGTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTGTGACCA 1719
QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAACGGAAGCGGCCCGGACCGCCCTACTGCTG 386
Db 1720 GGGCTGGGGCCCTATCAGTTATGCCAAACGGAAGCGGCCCGGACCGCCCTACTGCTG 1779
QY 387 GCACTACCCCCCAAAACCTTGGGTTATTTGCGCGGAAGAGTGTGTGGTCCGGTATA 446
Db 1780 GCACTACCCCCCAAAACCTTGGGTTATTTGCGCGGAAGAGTGTGTGGTCCGGTATA 1839
QY 447 TTGCTTCACTCCAGCCCGTGTGTGGGAACGACCGAGCTGGGGCGGCCACCTTA 506
Db 1840 TTGCTTCACTCCAGCCCGTGTGTGGGAACGACCGAGCTGGGGCGGCCACCTTA 1899
QY 507 CAGCTGGGTGAAATATGATACGAGACGTCTTCTGCTTAAACAATACGAGCCACCGTGGG 566
Db 1900 CAGCTGGGTGAAATATGATACGAGACGTCTTCTGCTTAAACAATACGAGCCACCGTGGG 1959
QY 567 CAATTGGTTGGTTGATCTGATGAACTCAACTGATGATTCACCAAGTGTGGGAGCGCC 626
Db 1960 CAATTGGTTGGTTGATCTGATGAACTCAACTGATGATTCACCAAGTGTGGGAGCGCC 2019
QY 627 TCTTGTGTCTATCGGAGGGCGGCAACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 686
Db 2020 TCTTGTGTCTATCGGAGGGCGGCAACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2079
QY 687 CAAGCATCCGGACCGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 746
Db 2080 CAAGCATCCGGACCGCCACATCTCTCGGTGCGGCTCCGGTCCCTGGATCACACCCAGGTG 2139
QY 747 CTTGCTCGACTACCCGTAAGGCTTTGGCAATTATCTTGTATCATCAATACCAATATT 806
Db 2140 CTTGCTCGACTACCCGTAAGGCTTTGGCAATTATCTTGTATCATCAATACCAATATT 2199
QY 807 TAAATCAGATGTAGCTGGGAGGGTCCGACACAGCTGGAGCTGCTGCACTGGAC 866
Db 2200 TAAATCAGATGTAGCTGGGAGGGTCCGACACAGCTGGAGCTGCTGCACTGGAC 2259
QY 867 GCGGGCGAAACGTTGGGATCTGGAAGATAGGGACAGGTCCGAGATC 912
Db 2260 GCGGGCGAAACGTTGGGATCTGGAAGATAGGGACAGGTCCGAGATC 2305
```

RESULT 7  
US-08-444-818-123/c  
; Sequence 123, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Rutter, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:

```
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 123:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ANTI-SENSE: YES
; US-08-444-818-123

Query Match 52.4%; Score 834.8; DB 3; Length 9185;
Best Local Similarity 96.4%; Pred. No. 2.3e-258;
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGTGCTGCTGTGTGGAGAGTCTTCTGTTTCGCCCCAGCGTAGCGAAACCCACGT 86
Db 7766 CTGGCGAAAGTCTCTGTAGTGTCTGTCTATTTGCGCGCTCGACGCGGAAACCCACGT 7707
QY 87 CACCGGGGAAGTGGCGGCCACACTGTGTCTGATTGTTAGCTTCTCGCACCGGCGC 146
Db 7706 CACCGGGGAAGTGGCGGCCACACTGTGTCTGATTGTTAGCTTCTCGCACCGGCGC 7647
QY 147 CAAGCAGAACGCTCCAGCTGATCAACACCAACGCGAGTTGGCACTCAATAGCAOGGCCCT 206
Db 7646 CAAGCAGAACGCTCCAGCTGATCAACACCAACGCGAGTTGGCACTCAATAGCAOGGCCCT 7587
QY 207 GAATCTCAATATAGCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCAAGTT 266
Db 7586 GAATCTCAATATAGCTCAACACCGGCTGTGTGGCAGGGCTTTTCTATCACCAAGTT 7527
QY 267 CAACTCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCGGACCCCTTACCGATTGTGACCA 326
Db 7526 CAACTCTTCAGGCTGTCTGAGAGGCTAGCCAGCTGCGGACCCCTTACCGATTGTGACCA 7467
QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAACGGAAGCGGCCCGGACCGCCCTACTGCTG 386
Db 7466 GGGCTGGGGCCCTATCAGTTATGCCAAACGGAAGCGGCCCGGACCGCCCTACTGCTG 7407
QY 387 GCACTACCCCCCAAAACCTTGGGTTATTTGTCGCCGGAAGAGTGTGTGGTCCGGTATA 446
Db 7406 GCACTACCCCCCAAAACCTTGGGTTATTTGTCGCCGGAAGAGTGTGTGGTCCGGTATA 7347
QY 447 TTGCTTCACTCCAGCCCGTGTGTGGGAACGACCGAGCTGGGGCGGCCACCTTA 506
Db 7346 TTGCTTCACTCCAGCCCGTGTGTGGGAACGACCGAGCTGGGGCGGCCACCTTA 7287
QY 507 CAGCTGGGTGAAATATGATACGGAAGTCTTCTGCTTAAACAATACGAGCCACCGTGGG 566
Db 507 CAGCTGGGTGAAATATGATACGGAAGTCTTCTGCTTAAACAATACGAGCCACCGTGGG 566
```







```

;
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4167_feature
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Leu."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4323_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4701_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Tyr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4752
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ser."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5970
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6183
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be His."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6186_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Cys."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6402_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7386_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ser."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7494_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Phe."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7497_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ala."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7845_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Phe."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8409_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9102_
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also by Gly."
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: 9327
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Pro."
US-07-910-760-9

Query Match 52.4%; Score 834.8; DB 1; Length 9401;
Best Local Similarity 96.4%; Pred. No. 2.3e-258;
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCGAGGCTAGGAAACCCACGT 86
DB 1442 CTGGGCGAAGGTCCTGGTAGTCTGCTGCTATTTCGCCGCGTCGACGCGGAAACCCACGT 1501
QY 87 CACCGGGGGAAGTGGCGGCCACACTGTGTCTGCTGATTGTGTAGCTCTCTCGCACGAGCGC 146
DB 1502 CACCGGGGGAAGTGGCGGCCACACTGTGTCTGCTGATTGTGTAGCTCTCTCGCACGAGCGC 1561
QY 147 CAAGCAGAACGTCCTCAGCTGATCAACACCAACGCGAGTGTGGCACCTCAATAGCACGCCCT 206
DB 1562 CAAGCAGAACGTCCTCAGCTGATCAACACCAACGCGAGTGTGGCACCTCAATAGCACGCCCT 1621
QY 207 GAACTGCAATGATAGCTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACAAGTT 266
DB 1622 GAACTGCAATGATAGCTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACAAGTT 1681
QY 267 CAACTCTTCAGGCTGTCTCTGAGAGGCTAGCCAGCTGCCGACCCCTTTACCGATTTCACCA 326
DB 1682 CAACTCTTCAGGCTGTCTCTGAGAGGCTAGCCAGCTGCCGACCCCTTTACCGATTTCACCA 1741
QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCGCCGACGAGCGCCCTTACTGTG 386
DB 1742 GGGCTGGGGCCCTATCAGTTATGCCAAGGAGCGGCCGCCGACGAGCGCCCTTACTGTG 1801
QY 387 GCACTACCCCCCAAAACCTTCGGGTATTGTGCCCGGGAAGAGTGTGTGTCGGGTATA 446
DB 1802 GCACTACCCCCCAAAACCTTCGGGTATTGTGCCCGGGAAGAGTGTGTGTCGGGTATA 1861
QY 447 TTGCTTCACTCCGACGCCCGTGGTGGGAACGACCGACAGTGGGGCGGCCACCTA 506
DB 1862 TTGCTTCACTCCGACGCCCGTGGTGGGAACGACCGACAGTGGGGCGGCCACCTA 1921
QY 507 CAGCTGGGGTGAAATGATACGACGCTCTTCGCTCTTAAACAATACAGGCCACCGCTGG 566
DB 1922 CAGCTGGGGTGAAATGATACGACGCTCTTCGCTCTTAAACAATACAGGCCACCGCTGG 1981
QY 567 CAAATTGGTTCCGTTTCTACTGTAACCTCAACTGGAATCACAAGTGTGCGGAGCGCC 626
DB 1982 CAAATTGGTTCCGTTTCTACTGTAACCTCAACTGGAATCACAAGTGTGCGGAGCGCC 2041
QY 627 TCCTTTGTCTATCGGAGGGGGGCAACAACACCTGCACTGCCACCTGATTGTTCG 686
DB 2042 TCCTTTGTCTATCGGAGGGGGGCAACAACACCTGCACTGCCACCTGATTGTTCG 2101
QY 687 CAAAGCATCCGAGCGCCACATCTCTCGTGGGGTCCGGTCCCTGGATCACAACCCAGGTG 746
DB 2102 CAAAGCATCCGAGCGCCACATCTCTCGTGGGGTCCGGTCCCTGGATCACAACCCAGGTG 2161
QY 747 CCTGTCGACTACCCGTATAGCTTTGGCATTTATCTTTGTACCATCACTACCATATT 806
DB 2162 CCTGTCGACTACCCGTATAGCTTTGGCATTTATCTTTGTACCATCACTACCATATT 2221
QY 807 TAAATCAGGATGTACGTGGGAGGGGTGGAACACAGGCTGGAAGTCCCTGCACTGGAC 866
DB 2222 TAAATCAGGATGTACGTGGGAGGGGTGGAACACAGGCTGGAAGTCCCTGCACTGGAC 2281
QY 867 GCGGGGCAACGTTGCGATCTGGAGATAGGGAAGAGTGGGACAGGTCCGAGATC 912
DB 2282 GCGGGGCAACGTTGCGATCTGGAGATAGGGAAGAGTGGGACAGGTCCGAGATC 2327

```

```

RESULT 11
US-08-440-519-9
; Sequence 9, Application US/08440519

```

Patent No. 5712087  
GENERAL INFORMATION:  
APPLICANT: Houghton, Michael  
APPLICANT: Choo, Qui-Lim  
APPLICANT: Kuo, George  
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: P.O. Box 8097 (Int. Prop. R-440)  
CITY: Emeryville  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94662-8097  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,519  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/910,760  
FILING DATE: 07-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Blackburn Esq., Robert P.  
REGISTRATION NUMBER: 30,447  
REFERENCE/DOCKET NUMBER: 0101.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2702  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9401 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 342..9374  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 366  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Arg."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 372  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Thr."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 867  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Thr."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1341  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Val."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2148  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ile."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2883  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Phe."  
OTHER INFORMATION: can also be Asn."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3681  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ser."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3690  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Thr."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 4167  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Leu."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 4323  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Val."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 4701  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Tyr."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 4752  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ser."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 5970  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Gly."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6183  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be His."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6186  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Cys."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 6402  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Val."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7386  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ser."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7494  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Phe."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7497  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Ala."  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 7845  
OTHER INFORMATION: /note= "This amino acid position  
OTHER INFORMATION: can also be Phe."  
FEATURE:  
NAME/KEY: misc\_feature

```

; LOCATION: 8409
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9102
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9327
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Pro."
; US-08-440-519-9

Query Match      52.4%; Score 834.8; DB 1; Length 9401;
Best Local Similarity 96.4%; Pred. No. 2.3e-258;
Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 27 CTGTGTCTCTCTGTGTGGAGCAGTCTTTCGTTTCGCCCGAGCGCTAGCGAAACCCACGTT 86
Db 1442 CTGGCGAAGTCTCTGTGTAGTGTCTGTCTATTTTCGGGGTGCACGCGAAACCCACGTT 1501

QY 87 CACCGGGGAAGTGC CGGCGCACACTGTGTCTGTGATTTGTAGCTCTCTGCACACGCGC 146
Db 1502 CACCGGGGAAGTGC CGGCGCACACTGTGTCTGTGATTTGTAGCTCTCTGCACACGCGC 1561

QY 147 CACGAGAGAGTCCAGTGTATCAACACCAACGCGAGTGTGCACCTCAATAGCAGCGCCT 206
Db 1562 CAAGCAGAGAGTCCAGTGTATCAACACCAACGCGAGTGTGCACCTCAATAGCAGCGCCT 1621

QY 207 GAAGTCAATGATAGTCTCAACACCGGCTGTTTGGCAGGCGCTTTCTATCACCACAAAGTT 266
Db 1622 GAAGTCAATGATAGTCTCAACACCGGCTGTTTGGCAGGCGCTTTCTATCACCACAAAGTT 1681

QY 267 CAAGTCTTCAAGGCTGTCTGAGAGGTAGCAGCTCGGACCCCTTACCGATTTTGACCA 326
Db 1682 CAAGTCTTCAAGGCTGTCTGAGAGGTAGCAGCTCGGACCCCTTACCGATTTTGACCA 1741

QY 327 GGGCTGGGCGCCCTATCAGTTATGCCAAGCGAGCGGCGCCCGACCGGCGCCCTACTGCTG 386
Db 1742 GGGCTGGGCGCCCTATCAGTTATGCCAAGCGAGCGGCGCCCGACCGGCGCCCTACTGCTG 1801

QY 387 GCACACTCCCCCAAAACCTTGCCTGATTTGTGCCGCGAAGAGTGTGTGTGTCGGGTATA 446
Db 1802 GCACACTCCCCCAAAACCTTGCCTGATTTGTGCCGCGAAGAGTGTGTGTGTCGGGTATA 1861

QY 447 TTGCTTCACTCCAGCGCCCGTGTGTGGAGCGACCGACAGTTCGGGCGCGCCACCTTA 506
Db 1862 TTGCTTCACTCCAGCGCCCGTGTGTGGAGCGACCGACAGTTCGGGCGCGCCACCTTA 1921

QY 507 CAGCTGGGTGAAATGATACGAGCTCTTCGTCCTTACAAATACAGGCGCCCGCTGGG 566
Db 1922 CAGCTGGGTGAAATGATACGAGCTCTTCGTCCTTACAAATACAGGCGCCCGCTGGG 1981

QY 567 CAATGCTGTCGTTGTACCTGTGATGAATCAACTGATTTCAACAAAGTGTGGAGCGCC 626
Db 1982 CAATGCTGTCGTTGTACCTGTGATGAATCAACTGATTTCAACAAAGTGTGGAGCGCC 2041

QY 627 TCCTTGTGTATCGAGGCGCGGCAACCAACCCCTGCACTGCCCGCACTGATTTGCTTCG 686
Db 2042 TCCTTGTGTATCGAGGCGCGGCAACCAACCCCTGCACTGCCCGCACTGATTTGCTTCG 2101

QY 687 CAAGCATCGGAGCGGCACATCTCTGGTGGCGCTCCGCTCCGTGATCAGACCCAGGTG 746
Db 2102 CAAGCATCGGAGCGGCACATCTCTGGTGGCGCTCCGCTCCGTGATCAGACCCAGGTG 2161

QY 747 CTGTGTCGACTACCCGTATAGGCTTTGGCATTTATCTTACCATCACTACACCATATT 806
Db 2162 CTGTGTCGACTACCCGTATAGGCTTTGGCATTTATCTTACCATCACTACACCATATT 2221

QY 807 TAAATACAGGATGACGTGGAGGGGTGGAACACAGGCTGGAAGCTGCTGCAACTGGAC 866
Db 807 TAAATACAGGATGACGTGGAGGGGTGGAACACAGGCTGGAAGCTGCTGCAACTGGAC 866

```

```

Db 2222 TAAATACAGGATGACGTGGAGGGGTGGAACACAGGCTGGAAGCTGCTGCAACTGGAC 2281
QY 867 GCGGGCGGAACGTTGCGATCTGGAACATAGGAGACAGTCCGAGATC 912
Db 2282 GCGGGCGGAACGTTGCGATCTGGAACATAGGAGACAGTCCGAGATC 2327

RESULT 12
US-08-440-549-9
; Sequence 9, Application US/08440549
; Patent No. 6312889
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,549
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Bq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 342..9374
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 366
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Arg."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 372
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 867
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1341
; OTHER INFORMATION: /note= "This amino acid position

```

```
;
; OTHER INFORMATION: can also be Val."
; FEATURE: misc_feature
; NAME/KEY: 2148
; LOCATION: 2148
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ile."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2883
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Asn."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3681
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ser."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3690
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Thr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4167
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Leu."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4323
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4701
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Tyr."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4752
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ser."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5970
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6183
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be His."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6186
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Cys."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6402
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Val."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7386
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ser."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7494
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Phe."
; FEATURE:
; NAME/KEY: misc_feature
```

```
;
; LOCATION: 7497
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Ala."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7845
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Phe."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8409
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9102
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also by Gly."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9327
; OTHER INFORMATION: /note= "This amino acid position
; OTHER INFORMATION: can also be Pro."
;
; US-08-440-549-9
;
; Query Match 52.4%; Score 834.8; DB 3; Length 9401;
; Best Local Similarity 96.4%; Pred. No. 2.3e-258;
; Matches 854; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
;
; QY 27 CTGTGTGCTGCTGTGTGTGAGGAGCTTTCGTTTCGCCCGAGCGTACGCGAAACCCACGT 86
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1442 CTGGCGGAAGGTCCTGTGTGTGTCTATTTTTCGGCGGCTGCGACGCGGAAACCCACGT 1501
;
; QY 87 CACGGGGGAAGTCCGGCCACACTGTCTCTGGATTTCCTAGCTCTCTCGCACCAAGGCGC 146
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1502 CACGGGGGAAGTCCGGCCACACTGTCTCTGGATTTCCTAGCTCTCTCGCACCAAGGCGC 1561
;
; QY 147 CAAGCAGAACGTCCTGAGTGTATCAACCAACGCGAGTTGGCAGCTCTCAATAGCAGGCGCT 206
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1562 CAAGCAGAACGTCCTGAGTGTATCAACCAACGCGAGTTGGCAGCTCTCAATAGCAGGCGCT 1621
;
; QY 207 GAATGCAATGATAGCTCAACACCGGCTGTTCGAGGGCTTTTCTATCACCACAAGTT 266
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1622 GAATGCAATGATAGCTCAACACCGGCTGTTCGAGGGCTTTTCTATCACCACAAGTT 1681
;
; QY 267 CAATCTTTGAGGCTGTCTGAGAGCTAGCAGCTGCGAGCCCTTACCGATTTTGACCA 326
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1682 CAATCTTTGAGGCTGTCTGAGAGCTAGCAGCTGCGAGCCCTTACCGATTTTGACCA 1741
;
; QY 327 GGGCTGGGGCCCTATCAGTTATGCGAAGCGGCGCCCGCACAGCGGCCCTACTGCTG 386
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1742 GGGCTGGGGCCCTATCAGTTATGCGAAGCGGCGCCCGCACAGCGGCCCTACTGCTG 1801
;
; QY 387 GCATCTCCCGCCCAAAACCTTGGGTATTTGCGCCGGAAGAGTGTTGGTCCGCTATA 446
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1802 GCATCTCCCGCCCAAAACCTTGGGTATTTGCGCCGGAAGAGTGTTGGTCCGCTATA 1861
;
; QY 447 TTGCTTCACTCCAGCCCGCTGTGTGTGGGAACGACAGCTGCGGCGCGCCACCTTA 506
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1862 TTGCTTCACTCCAGCCCGCTGTGTGTGGGAACGACAGCTGCGGCGCGCCACCTTA 1921
;
; QY 507 CAGCTGGGGTGAATGATAGGAGCTTTCGTCCTTAACAATACAGGCGCAACCGTGGG 566
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1922 CAGCTGGGGTGAATGATAGGAGCTTTCGTCCTTAACAATACAGGCGCAACCGTGGG 1981
;
; QY 567 CAATTTGTTGGTTTCTGCTGATGAACCTCAACTGATTTCACCAAGTGTGGGAGCGGC 626
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 1982 CAATTTGTTGGTTTCTGCTGATGAACCTCAACTGATTTCACCAAGTGTGGGAGCGGC 2041
;
; QY 627 TCCTTGTGTGATCGGAGGGCGGGCAACAACCTTGCACTGCCCACTGATTGCTTCGG 686
; | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2042 TCCTTGTGTGATCGGAGGGCGGGCAACAACCTTGCACTGCCCACTGATTGCTTCGG 2101
```







[illegible]

Search completed: April 15, 2005, 02:48:19  
Job time : 376 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 23:14:12 ; Search time 964 Seconds  
(without alignments)  
10025.177 Million cell updates/sec

Title: US-10-715-665-6\_COPY\_1992\_3584

Perfect score: 1593

Sequence: 1 atgtgatgaatgaagagag.....tttgtctctgggtatacatt 1593

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

#### Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	836.4	52.5	1989	17 US-10-371-040-3	Sequence 3, Appli
2	834.8	52.4	9379	9 US-09-916-359-1	Sequence 1, Appli
3	834.8	52.4	9401	18 US-10-445-724-1	Sequence 1, Appli
4	833.2	52.3	1914	15 US-10-187-257-3	Sequence 3, Appli
5	833.2	52.3	1914	16 US-10-265-083-1	Sequence 1, Appli
6	745.2	46.8	9646	9 US-09-742-659-3	Sequence 3, Appli
7	745.2	46.8	9646	9 US-09-238-076-1	Sequence 1, Appli
8	745.2	46.8	9646	10 US-09-995-937-1	Sequence 1, Appli
9	745.2	46.8	9646	10 US-09-917-563-1	Sequence 1, Appli
10	745.2	46.8	12980	9 US-09-238-076-5	Sequence 5, Appli
11	745.2	46.8	12980	10 US-09-995-937-5	Sequence 5, Appli

12	745.2	46.8	12980	10	US-09-917-563-5	Sequence 5, Appli
13	742	46.6	1845	17	US-10-365-620-73	Sequence 73, Appli
14	742	46.6	1845	19	US-10-912-969-75	Sequence 75, Appli
15	742	46.6	2517	17	US-10-365-620-75	Sequence 75, Appli
16	742	46.6	2517	19	US-10-912-969-77	Sequence 77, Appli
17	742	46.6	2517	19	US-10-913-171-48	Sequence 48, Appli
18	742	46.6	9599	17	US-10-189-359-13	Sequence 13, Appli
19	742	46.6	10803	9	US-09-747-419-17	Sequence 17, Appli
20	742	46.6	10803	15	US-10-259-275-17	Sequence 17, Appli
21	740.4	46.5	9416	9	US-09-929-955-13	Sequence 13, Appli
22	740.4	46.5	9416	13	US-10-104-968-13	Sequence 13, Appli
23	740.4	46.5	9416	17	US-10-719-619-13	Sequence 13, Appli
24	740.4	46.5	9416	18	US-10-817-591-13	Sequence 13, Appli
25	739.4	46.4	9365	10	US-09-827-688-7	Sequence 7, Appli
26	738.8	46.4	9416	9	US-09-238-076-19	Sequence 19, Appli
27	738.8	46.4	9416	10	US-09-995-937-19	Sequence 19, Appli
28	738.8	46.4	9416	10	US-09-917-563-19	Sequence 19, Appli
29	738.8	46.4	9622	18	US-10-475-989-2	Sequence 2, Appli
30	738	46.3	1266	19	US-10-912-969-71	Sequence 71, Appli
31	738	46.3	1938	17	US-10-365-620-71	Sequence 71, Appli
32	738	46.3	1938	19	US-10-912-969-73	Sequence 73, Appli
33	738	46.3	1938	19	US-10-913-171-46	Sequence 46, Appli
34	679.4	42.6	4084	10	US-09-781-891D-21	Sequence 21, Appli
35	679.4	42.6	4496	10	US-09-781-891D-22	Sequence 22, Appli
36	678	42.6	1696	15	US-10-267-922-15	Sequence 15, Appli
37	676.4	42.5	750	9	US-09-952-572-1	Sequence 1, Appli
38	676.2	42.4	1365	17	US-10-365-620-33	Sequence 33, Appli
39	676.2	42.4	1365	19	US-10-912-969-37	Sequence 37, Appli
40	676.2	42.4	2037	17	US-10-365-620-28	Sequence 28, Appli
41	676.2	42.4	2037	17	US-10-365-620-31	Sequence 31, Appli
42	676.2	42.4	2037	19	US-10-912-969-35	Sequence 35, Appli
43	676.2	42.4	2037	19	US-10-913-171-28	Sequence 28, Appli
44	674.8	42.4	3200	17	US-10-453-792-280	Sequence 280, App
45	674.8	42.4	3221	19	US-10-856-355-29	Sequence 29, Appli

#### ALIGNMENTS

RESULT 1  
US-10-371-040-3  
; Sequence 3, Application US/10371040  
; Publication No. US20040001854A1  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Abirmani, Sergio  
; APPLICANT: Chien, David  
; APPLICANT: Selby, Mark  
; TITLE OF INVENTION: Intracellular Production of Hepatitis C E1 and E2  
; TITLE OF INVENTION: Truncated Polypeptides  
; FILE REFERENCE: 1378.002  
; CURRENT APPLICATION NUMBER: US/10/371,040  
; PRIOR FILING DATE: 2003-02-18  
; PRIOR APPLICATION NUMBER: US/09/073,406  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/045,675  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-06  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1989  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1989)  
US-10-371-040-3

Query Match 52.5%; Score 836.4; DB 17; Length 1989;  
Best Local Similarity 96.5%; Pred. No. 9.2e+268;  
Matches 855; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy	27	CTGTGTGCTGCTGCTGTGTGGAGCAGTCTTTCGTTTCGCCCAGCGCTAGCGAAACCCACGT	86
Db	12		
		CTGGCGAAGGTCCTGGTAGTGTCTGCTATTTTTCGGCGTGCAGCGGAAACCCACGT	71
Qy	87	CACCGGGGAAGTGGCGGCGCACATGTGTCTGGAATTTGTAGCCTCTCTGCACACAGCGC	146
Db	72		
		CACCGGGGAAGTGGCGGCGCACATGTGTCTGGAATTTGTAGCCTCTCTGCACACAGCGC	131
Qy	147	CAAGCAGACGTCACGCTGATCAACACCAACGGCAGTTGGCAGCTCAATAGCAGCGCCT	206
Db	132		
		CAAGCAGACGTCACGCTGATCAACCAACGGCAGTTGGCAGCTCAATAGCAGCGCCT	191
Qy	207	GAACTGCAATGATAGCCTCAAACACGGGCTGGTGTGGCAGGGCTTTTCTATCACCACAAGTT	266
Db	192		
		GAACTGCAATGATAGCCTCAAACACGGGCTGGTGTGGCAGGGCTTTTCTATCACCACAAGTT	251
Qy	267	CAACTCTTCAGGCTGTCTGAGAGCTAGCAGCTGCGACCCCTTACCGAATTTTGACCA	326
Db	252		
		CAACTCTTCAGGCTGTCTGAGAGCTAGCAGCTGCGACCCCTTACCGAATTTTGACCA	311
Qy	327	GGGCTGGGGCCCTATCAGTTATGCCAACGGAAAGCGGCCCGACAGCGCCCTTACTGCTG	386
Db	312		
		GGGCTGGGGCCCTATCAGTTATGCCAACGGAAAGCGGCCCGACAGCGCCCTTACTGCTG	371
Qy	387	GCACTACCCCCCAAAACCTTCGGGTATTTGTGCCCGGAAAGTGTGTGGTCCGGTATA	446
Db	372		
		GCACTACCCCCCAAAACCTTCGGGTATTTGTGCCCGGAAAGTGTGTGGTCCGGTATA	431
Qy	447	TTGCTTTCATCCCGCCCGTGGTGGGAACGACCGACAGGTCGGGCGCGCCACCTA	506
Db	432		
		TTGCTTTCATCCCGCCCGTGGTGGGAACGACCGACAGGTCGGGCGCGCCACCTA	491
Qy	507	CAGTCGGGTGAAATGATACGGAAGTCTTCGTCTTAAACAATACGAGCCACGCTGGG	566
Db	492		
		CAGTCGGGTGAAATGATACGGAAGTCTTCGTCTTAAACAATACGAGCCACGCTGGG	551
Qy	567	CAATTGGTTGGTGTGTAACCTGGAATGAACCTGAATTCACAAAGTGTGGGAGCGCC	626
Db	552		
		CAATTGGTTGGTGTGTAACCTGGAATGAACCTGAATTCACAAAGTGTGGGAGCGCC	611
Qy	627	TCCCTTGTCATCGGAGGGCGGGCAACAACCTGCAGCTGCCCACTGATTTGCTTCGG	686
Db	612		
		TCCCTTGTCATCGGAGGGCGGGCAACAACCTGCAGCTGCCCACTGATTTGCTTCGG	671
Qy	687	CAAGCATCCGGAGCGCACATCTCTCGTGGGGCTCCGGTCCCTGGATCAACCCAGGTG	746
Db	672		
		CAAGCATCCGGAGCGCACATCTCTCGTGGGGCTCCGGTCCCTGGATCAACCCAGGTG	731
Qy	747	CCTGTCGACTACCCGTAATAGGCTTTGGATTAATCTTGTACCAATCAACTACACCATATT	806
Db	732		
		CCTGTCGACTACCCGTAATAGGCTTTGGCATTAATCTTGTACCAATCAACTACACCATATT	791
Qy	807	TAAATACAGATGACTCGTGGAGGGGTGGAACACAGGCTGGAAGCTGCCTGCAACTGGAC	866
Db	792		
		TAAATACAGATGACTCGTGGAGGGGTGGAACACAGGCTGGAAGCTGCCTGCAACTGGAC	851
Qy	867	GGCGGGCGAACGTTGCGATCTCGAAGATAGGGACAGGTCGAGATC	912
Db	852	GGCGGGCGAACGTTGCGATCTCGAAGATAGGGACAGGTCGAGCTC	897

## RESULT 2

US-09-916-359-1

03-03-210-333-1  
; Sequence 1, Application US/09916359

; Patent No. US20020034734A1

FACEIT NO: 052002003  
; GENERAL INFORMATION:

APPLICANT: Veronique Barban

;  
; AFFILIANT: veronique baldan  
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR

**TITLE OF INVENTION:** VACCINE COMPOSITION FOR TREATING C HEPATITIS

; TITLE OF INVENTION: TREATING C REPAIRS  
 : FILE REFERENCE: PMCF97-03A

FILE REFERENCE: FMCF97-03A  
CURRENT APPLICATION NUMBER: IIS/0



```

QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAACGGAAGCGCCCGGACCGCCCTACTGCTG 386
DB 885 GGGCTGGGGCCCTATCAGTTATGCCAAACGGAAGCGCCCGGACCGCCCTACTGCTG 944
QY 387 GCACTACCCGCCAAAACCTTGGGTTATGTGCCCGGGAAGAGTGTGTGGTCCGGTATA 446
DB 945 GCACTACCCGCCAAAACCTTGGGTTATGTGCCCGGGAAGAGTGTGTGGTCCGGTATA 1004
QY 447 TTGCTTCACTCCAGCGCCCGTGTGTGGAAACGACCGACAGTCGGGCGGCCACCTTA 506
DB 1005 TTGCTTCACTCCAGCGCCCGTGTGTGGAAACGACCGACAGTCGGGCGGCCACCTTA 1064
QY 507 CAGCTGGGTGAAATGATACGAGCTCTTGGTCCCTTACAAATACACGCGCCACCGCTGG 566
DB 1065 CAGCTGGGTGAAATGATACGAGCTCTTGGTCCCTTACAAATACACGCGCCACCGCTGG 1124
QY 567 CAATTGGTTCCGTTGACCTGGATGAATCACTCACTGGATTCAACAAAGTGTGGGAGCGCC 626
DB 1125 CAATTGGTTCCGTTGACCTGGATGAATCACTCACTGGATTCAACAAAGTGTGGGAGCGCC 1184
QY 627 TCCTTGTGTATCGGAGGGCGGGCAACAAACCCCTGCACTGCCCCCACTGATTGCTTCG 686
DB 1185 TCCTTGTGTATCGGAGGGCGGGCAACAAACCCCTGCACTGCCCCCACTGATTGCTTCG 1244
QY 687 CAAGCATCCGAGCGCCACATATCTCTCGGTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 746
DB 1245 CAAGCATCCGAGCGCCACATATCTCTCGGTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 1304
QY 747 CTGTGTGCACTACCCGTTAGGCTTTGGCAATTAATCTTGTACCATCAACTACACCATATT 806
DB 1305 CTGTGTGCACTACCCGTTAGGCTTTGGCAATTAATCTTGTACCATCAACTACACCATATT 1364
QY 807 TAAATCAGGATGATCGTGGGAGGGTCGAACACAGCTGGAGCTGCCTGCAACTGGAC 866
DB 1365 TAAATCAGGATGATCGTGGGAGGGTCGAACACAGCTGGAGCTGCCTGCAACTGGAC 1424
QY 867 GCGGGCGGACGTTGGATCTGGAAGATAGGACAGGTCGGAGATC 912
DB 1425 GCGGGCGGACGTTGGATCTGGAAGATAGGACAGGTCGGAGCTC 1470

```

```

RESULT 5
US-10-265-083-1
; Sequence 1, Application US/10265083
; Publication No. US20030170273A1
; GENERAL INFORMATION:
; APPLICANT: O'HAGAN, Derek
; APPLICANT: VALIANTE, Nicholas
; TITLE OF INVENTION: ADJUVANT COMPOSITIONS
; FILE REFERENCE: 2302-18357.30
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1911)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: HCV-1 E1/E2/p7
; OTHER INFORMATION: region
US-10-265-083-1

```

```

Query Match 52.3%; Score 833.2; DB 16; Length 1914;
Best Local Similarity 96.3%; Pred. No. 1.1e-266;
Matches 853; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 27 CTGTGTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCGAGCGCTAGCGAAACCCACGT 86
DB 585 CTGGCGGAAGGTCCTGTGTAGTGTGCTGCTGCTATTTGCGGGGTGACGCGGAAACCCACGT 644

```

```

QY 87 CACCGGGGAGTGGCGGCCACATGTGTCTGGATTTGTAGCTCTCTCGCACCGAGCGC 146
DB 645 CACCGGGGAGTGGCGGCCACATGTGTCTGGATTTGTAGCTCTCTCGCACCGAGCGC 704
QY 147 CAAGCAGAACGTCAGCTGATCAACACCAACGGCAGTTGGCACCTCAATAGCACGCGCCT 206
DB 705 CAAGCAGAACGTCAGCTGATCAACACCAACGGCAGTTGGCACCTCAATAGCACGCGCCT 764
QY 207 GAACTGCAATGATAGCTTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACAAGTT 266
DB 765 GAACTGCAATGATAGCTTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACAAGTT 824
QY 267 GAACTGCAATGATAGCTTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACAAGTT 326
DB 825 GAACTGCAATGATAGCTTCAACACCGGCTGGTGGCAGGGCTTTTCTATCACCACAAGTT 884
QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAACGGAGCGGCCCGACCGAGCGCCCTACTGCTG 386
DB 885 GGGCTGGGGCCCTATCAGTTATGCCAAACGGAGCGGCCCGACCGAGCGCCCTACTGCTG 944
QY 387 GCACTACCCGCCAAAACCTTGGGTTATTTGTGCCCGGGAAGAGTGTGTGGTCCGGTATA 446
DB 945 GCACTACCCGCCAAAACCTTGGGTTATTTGTGCCCGGGAAGAGTGTGTGGTCCGGTATA 1004
QY 447 TTGCTTCACTCCAGCGCCCGTGTGGTGGACCGACCGACAGCTCGGGCGGCCCCACCTTA 506
DB 1005 TTGCTTCACTCCAGCGCCCGTGTGGTGGACCGACCGACAGCTCGGGCGGCCCCACCTTA 1064
QY 507 CAGCTGGGGTGAATATGATACGAGCTCTTTCGCTCTTAACAAATACAGGCCACCGCTGG 566
DB 1065 CAGCTGGGGTGAATATGATACGAGCTCTTTCGCTCTTAACAAATACAGGCCACCGCTGG 1124
QY 567 CAATTGGTTCCGTTGATCTGGATGAACTCAACTGGATTCAACAAAGTGTGGGAGCGCC 626
DB 1125 CAATTGGTTCCGTTGATCTGGATGAACTCAACTGGATTCAACAAAGTGTGGGAGCGCC 1184
QY 627 TCCTTGTGTATCGGAGGGCGGGCAACAAACACCTTGCACTCCGCCACTGATTGCTTCG 686
DB 1185 TCCTTGTGTATCGGAGGGCGGGCAACAAACACCTTGCACTCCGCCACTGATTGCTTCG 1244
QY 687 CAAGCATCCGAGCGCCACATATCTCTCGGTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 746
DB 1245 CAAGCATCCGAGCGCCACATATCTCTCGGTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 1304
QY 747 CTGTGTGCACTACCCGTTAGGCTTTGGCAATTAATCTTGTACCATCAACTACACCATATT 806
DB 1305 CTGTGTGCACTACCCGTTAGGCTTTGGCAATTAATCTTGTACCATCAACTACACCATATT 1364
QY 807 TAAATCAGGATGATCGTGGGAGGGTCGAACACAGCTGGAGCTGCCTGCAACTGGAC 866
DB 1365 TAAATCAGGATGATCGTGGGAGGGTCGAACACAGCTGGAGCTGCCTGCAACTGGAC 1424
QY 867 GCGGGCGGAAACGTTGGATCTGGAAGATAGGACAGGTCGGAGATC 912
DB 1425 GCGGGCGGAAACGTTGGATCTGGAAGATAGGACAGGTCGGAGCTC 1470

```

```

RESULT 6
US-09-742-659-3
; Sequence 3, Application US/09742659
; Patent No. US20010034019A1
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallo, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659

```





QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAAGCGAGCGCCCGACAGCGCCCTACTGCTG 386  
 Db 1742 GGGCTGGGGTCCCTATCAGTTATGCCAAAGCGAGCGCCCTCGAGCAAGCGCCCTACTGCTG 1801  
 QY 387 GCACTACCCCGCCAAAACCTTGCGGTATGTGCCCGGAAAGAGTGTGTGGTCCGGTATA 446  
 Db 1802 GCACCTACCTCCAAAGACCTTGTCGATGTGTGCCCGCAAGAGCGTGTGTGGCCGGTATA 1861  
 QY 447 TTGCTTCACTCCAGCGCCCGTGTGTGGAAAGCAAGCGAGTGTGGGGCGGCCACCTTA 506  
 Db 1862 TTGCTTCACTCCAGCGCCCGTGTGTGGAAAGCAAGCGAGTGTGGGGCGGCCCTACCTTA 1921  
 QY 507 CAGCTGGGTGAAATGATACGAGCTCTTGTGCTTAAACAATACCAAGCGCCCGCTGGG 566  
 Db 1922 CAGCTGGGTGCAAAATGATACGAGTGTCTTGTGCTTAAACAATACCAAGCGCCCGCTGGG 1981  
 QY 567 CAATTGCTTGGGTGTACCTGTGATGAACCTCAACTGGAATTCACCAAGTGTGGAGCGCC 626  
 Db 1982 CAATTGCTTGGGTGTACCTGTGATGAACCTCAACTGGAATTCACCAAGTGTGGAGCGCC 2041  
 QY 627 TCCTTGTGATCGAGGGCGGGGCAACACACCTTGCACTGCGCCCACTGATTTGCTTCG 686  
 Db 2042 CCCTTGTGATCGAGGGGTGGGCAACACACCTTGTCTGTGCGCCCACTGATTTGCTTCG 2101  
 QY 687 CAGCATCCGAGCGGCACATACCTCTCGTCCGCTCGGTCCTGGATCACACCCAGGTG 746  
 Db 2102 CAAGCATCCGAGCGGCACATACCTCTCGTCCGCTCGGTCCTGGATCACACCCAGGTG 2161  
 QY 747 CTGCTCGACTACCCGTATAGGCTTTGGCATATCTTGTACCATCAACTACACCATATT 806  
 Db 2162 CATGGTCGACTACCCGTATAGGCTTTGGCATATCTTGTACCATCAACTACACCATATT 2221  
 QY 807 TAAATCAGATGTACGTGGAGGGGTGCAACACAGCTGGAGAGTGCCTGCACTGGAC 866  
 Db 2222 CAAAGTCAGATGTACGTGGAGGGGTGCAACACAGCTGGAGAGTGCCTGCACTGGAC 2281  
 QY 867 GCGGGCGAAGCTTGGCATCTGGAAGATAGGAGCAGCTCCGAGATC 912  
 Db 2282 GCGGGCGAAGCTTGGCATCTGGAAGATAGGAGCAGCTCCGAGATC 2327

RESULT 8

US-09-995-937-1  
 ; Sequence 1, Application US/09995937  
 ; Publication No. US20030028010A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RICE, CHARLES et al.  
 ; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
 ; STREET: 7733 FORSYTH BLVD., SUITE 1400  
 ; CITY: ST. LOUIS  
 ; STATE: MO  
 ; COUNTRY: USA  
 ; ZIP: 63105  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/995,937  
 ; FILING DATE: 28-No. US20030028010A1-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/034,756  
 ; FILING DATE: 04-May-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: HOLLAND, DONALD R.  
 ; REGISTRATION NUMBER: 35,197  
 ; REFERENCE/DOCKET NUMBER: 6029-4831

TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 314-727-5188  
 ; TELEFAX: 314-727-6092  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9646 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHEetical: NO  
 ; ANTI-SENSE: NO  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 ; US-09-995-937-1  
 ;  
 ; Query Match 46.8%; Score 745.2; DB 10; Length 9646;  
 ; Best Local Similarity 90.1%; Pred. No. 6.4e-237;  
 ; Matches 798; Conservative 0; Mismatches 88; Indels 0; Gaps 0;  
 ;  
 ; QY 27 CTGCTGCTGCTGCTGTGTGGAGCAGTCTTCGTTTCGCCCGAGCTAGCGAAACCCACGT 86  
 ; Db 1442 CTGGGCGAAGTCTCTGTGTAGTCTGCTATTTGCGGCGTCTGAGCGGAAACCCACGT 1501  
 ; QY 87 CACCGGGGGAAGTGTGCGGCCACACCTGTCTGTGGAATTTGTAGCTCTCTCGCACCGAGCGC 146  
 ; Db 1502 CACCGGGGGAAGTGTGCGGCCCGCACACCGCTGGCTTGTGTCTCTTACACCGAGCGC 1561  
 ; QY 147 CAAGCAGAAAGTGTGAGTGTATCAACCAACCGGAGTGTGGACCTCAATAGCAGCGCCCT 206  
 ; Db 1562 CAAGCAGAAAGTGTGAGTGTATCAACCAACCGGAGTGTGGACCTCAATAGCAGCGCCCT 1621  
 ; QY 207 GAATGCAATGATAGCTCAACACCGGCTGTGTGGAGGCTTTTCTATCACCAAGTT 266  
 ; Db 1622 GAATGCAATGAAAGCTTTAACCAGCTGTGTGTAGGAGCTCTTCTATCAGACCAAT 1681  
 ; QY 267 CAACTCTTCAGGCTGTCTGAGAGGCTAGCAGCTGTGCGACCCCTTACCGATTTTGACCA 326  
 ; Db 1682 CAACTCTTCAGGCTGTCTGAGAGGTTGGCCAGCTGTGCGACGCTTACCGATTTTGCCCA 1741  
 ; QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAAGCGAGCGCCCGACAGCGCCCTACTGCTG 386  
 ; Db 1742 GGGCTGGGGTCTCTATCAGTTATGCCAAAGCGAGCGCTTCGACAAACCGCCCTACTGCTG 1801  
 ; QY 387 GCACTACCCCGCCAAAACCTTGCGGTATTTGTCGCCGGAAGAGTGTGTGTGTCGGTATA 446  
 ; Db 1802 GCACTACCTCCAAAGACCTTGTCGATTTGTGCCCGCAAGAGCGTGTGTGGCCGGTATA 1861  
 ; QY 447 TTGCTTCACTCCAGCGCCCGTGTGTGGAAAGCAAGCGAGTGTGGGGCGGCCACCTTA 506  
 ; Db 1862 TTGCTTCACTCCAGCGCCCGTGTGTGGAAAGCAAGCGAGTGTGGGGCGGCCCTACCTTA 1921  
 ; QY 507 CAGCTGGGTGAAATGATACGAGCTCTTGTGCTCTTAAACAATACCAAGCGCCCGCTGGG 566  
 ; Db 1922 CAGCTGGGTGCAAAATGATACGAGTGTCTTGTGCTTAAACAATACCAAGCGCCCGCTGGG 1981  
 ; QY 567 CAATTGCTTGGGTGTACCTGTGATGAACCTCAACTGGAATTCACCAAGTGTGGAGCGCC 626  
 ; Db 1982 CAATTGCTTGGGTGTACCTGTGATGAACCTCAACTGGAATTCACCAAGTGTGGAGCGCC 2041  
 ; QY 627 TCCTTGTGATCGAGGGCGGGGCAACACACCTTGCACTGCGCCCACTGATTTGCTTCG 686  
 ; Db 2042 CCCTTGTGATCGAGGGGTGGGCAACACACCTTGTCTGTGCGCCCACTGATTTGCTTCG 2101  
 ; QY 687 CAGCATCCGAGCGGCACATACCTCTCGTCCGCTCGGTCCTGGATCACACCCAGGTG 746  
 ; Db 2102 CAAGCATCCGAGCGGCACATACCTCTCGTCCGCTCGGTCCTGGATCACACCCAGGTG 2161  
 ; QY 747 CCGTGTGCACTACCCGTATAGGCTTTGGCATATCTTGTACCATCAACTACACCATATT 806  
 ; Db 2162 CATGGTCGACTACCCGTATAGGCTTTGGCATATCTTGTACCATCAACTACACCATATT 2221  
 ; QY 807 TAAATCAGATGTACGTGGAGGGGTGCAACACAGCTGGAGTGCCTGCACTGGAC 866

Db 2222 CAAAGTCAGGATGTCGTGGAGGGTCGAGCAGCAGGCTGGAAAGCGCCTGCAATGGAC 2281  
 QY 867 GCGGGCGGAGCGTTGGATCTGGAAGATAGGAGCAGGTCGAGATC 912  
 Db 2282 GCGGGCGGAGCGTTGGATCTGGAAGACAGGACAGGTCGAGCTC 2327

RESULT 9

US-09-917-563-1  
 ; Sequence 1, Application US/09917563  
 ; Publication No. US20030073080A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RICE, CHARLES et al.  
 ; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C  
 ; VIRUS (HCV) AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
 ; STREET: 7733 FORSYTH BLVD., SUITE 1400  
 ; CITY: ST. LOUIS  
 ; STATE: MO  
 ; COUNTRY: USA  
 ; ZIP: 63105  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION NUMBER: US/09/917,563  
 ; FILING DATE: 27-Jul-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/238,076  
 ; FILING DATE: 26-JAN-1999  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: HOLLAND, DONALD R.  
 ; REGISTRATION NUMBER: 35,197  
 ; REFERENCE/DOCKET NUMBER: 6029-4831  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 314-727-5188  
 ; TELEFAX: 314-727-6092  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 9646 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-917-563-1  
 Query Match 46.8%; Score 745.2; DB 10; Length 9646;  
 Best Local Similarity 90.1%; Pred. No. 6.4e-237;  
 Matches 798; Conservative 0; Mismatches 88; Indels 0; Gaps 0;  
 QY 27 CTGTGCTGCTGCTGTGGAGCAGTCTTCGTCGCCCGGCTAGCGAAACCCAGT 86  
 Db 1442 CTGGGCGAAGTCTTGGTGTGCTGTCTATTTGCCGCGCTGACGGGAAACCCAGT 1501  
 QY 87 CACCGGGGGAAGTGGCGGCGCACACTGTGTCTGGATTTGTAGCTCTCGCACACGAGGCG 146  
 Db 1502 CACCGGGGGAAGTGGCGGCGCACACCGCTGGGCTTGTGTCTCTTACACGAGGCG 1561  
 QY 147 CAAGCAGAACGTCAGCTGATCAACACCGGAGTTGGACCTCAATAGCAGCGCCCT 206  
 Db 1562 CAAGCAGAACATCCAACTGATCAACACCAACGCGAGTTGGACATCAATAGCAGCGCTT 1621  
 QY 207 GAATGCAATGATAGCTCAACACCGGCTGGTTGGCAGGGCTTTCTATCACCACAAGTT 266  
 Db 1622 GAATGCAATGAAGCCTTAAACCGGCTGGTTAGCAGGGCTCTTCTATCAGCACAAAT 1681

QY 267 CAACTCTTCAGGCTGTCTGAGAGGCTAGCCAGCTCCGACCCCTTACCGATTTTGCACCA 356  
 Db 1682 CAACTCTTCAGGCTGTCTGAGAGGTTGGCCAGCTGCCACGCTTACCGATTTTGCACCA 1741  
 QY 327 GGGCTGGGCGCTTATCAGTTATGCCAAACGGAAGCGGCCCGACAGCGCCCTTACTGCTG 386  
 Db 1742 GGGCTGGGCGCTTATCAGTTATGCCAAACGGAAGCGGCTCGACGAAACGCGCTTACTGCTG 1801  
 QY 387 GCACCTACCCCGCAAAACCTTGGGTTATTTGTCGCCCGGAGAGGTGTGTGTCGCGGTATA 446  
 Db 1802 GCACCTACCCCTCCAAAGACCTTGTGGCATTTGTGCCCGCAAGAGCGGTGTGTGTCGCGGTATA 1861  
 QY 447 TTGCTTCACTCCAGCGCCCGTGTGGGAAACGACCGACAGGTCCGGGCGCGCCACCTA 506  
 Db 1862 TTGCTTCACTCCAGCGCCCGTGTGGGAAACGACCGACAGGTCCGGGCGCGCTTACTA 1921  
 QY 507 CAGCTGGGCTGAAATGATACGAGCGTCTTCTGCTCTTAAACAATACAGGCGCACCGCTGGG 566  
 Db 1922 CAGCTGGGCTGAAATGATACGAGCGTCTTCTGCTCTTAAACAATACAGGCGCACCGCTGGG 1981  
 QY 567 CAATTGGTTCCGTTGTACTGATGAACCTCAACTGGATTCAACAAGTGTGCGGAGCGCC 626  
 Db 1982 CAATTGGTTCCGTTGTACTGATGAACCTCAACTGGATTCAACAAGTGTGCGGAGCGCC 2041  
 QY 627 TCCTTGTGTATCGGAGGGCGGCAACAAACCCCTGCACTGCCCTGCTGCTTTCCTG 686  
 Db 2042 CCCTTGTGTATCGGAGGGGTGGGCAACAACCTTGTCTGCTGCCCTGCTGCTTTCCTG 2101  
 QY 687 CAAGCATCCGAGCGCCACATCTCTCGGTGCGGCTCCGCTCCCTGGATCACACCCAGGTG 746  
 Db 2102 CAAGCATCCGAGCGCCACATCTCTCGGTGCGGCTCCGCTCCCTGGATCACACCCAGGTG 2161  
 QY 747 CCGTGTGCACTACCGGTATAGGCTTTGGCATTTATCTTGTACCATCAATTCACCATATT 806  
 Db 2162 CATGTCGACTACCGGTATAGGCTTTGGCATTTATCTTGTACCATCAATTCACCATATT 2221  
 QY 807 TAAATCAGGATGTACGTGGAGGGGTGAAACACAGGCTGGAAGCTCCCTGCACTGGAC 866  
 Db 2222 CAACTCAGGATGTACGTGGAGGGGTGAGCAGCAGGCTGGAGCGGCTGCACTGGAC 2281  
 QY 867 GCGGGCGGAAAGTTCGATCTGGAAGATAGGAGCAGGTCGAGATC 912  
 Db 2282 GCGGGCGGAAAGTTCGATCTGGAAGATAGGAGCAGGTCGAGATC 2327

RESULT 10

US-09-238-076-5  
 ; Sequence 5, Application US/09238076  
 ; Patent No. US20020102540A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: RICE, CHARLES et al.  
 ; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C  
 ; VIRUS (HCV) AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 21  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
 ; STREET: 7733 FORSYTH BLVD., SUITE 1400  
 ; CITY: ST. LOUIS  
 ; STATE: MO  
 ; COUNTRY: USA  
 ; ZIP: 63105  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/238,076  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 09/034,756

FILING DATE: 2162 CATGTCGACTACCGCTATAGGCTTTGGCACTATCTTGTACCATCAATTACACCATATT 2221  
ATTORNEY/AGENT INFORMATION: 807 TAAATCAGGATGTACGTGGAGGGTCAACACAGGCTGGAAGCTGCCTGCAACTGGAC 866  
NAME: HOLLAND, DONALD R. 2222 CAAAGTCAGGATGTACGTGGAGGGTCAACACAGGCTGGAAGCGGCTGCAACTGGAC 2281  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 6029-4831  
TELECOMMUNICATION INFORMATION: 867 GCGGGGCGAACTGTCGATCTGGAAGATAGGAGCAGGTCCGAGATC 912  
TELEPHONE: 314-727-5188 2282 GCGGGGCGAACTGTCGATCTGGAAGACAGGACAGGTCCGAGCTC 2327  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12980 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-238-076-5

Query Match 46.8%; Score 745.2; DB 9; Length 12980;  
Best Local Similarity 90.1%; Pred. No. 7.5e-237;  
Matches 798; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 27 CTGTGCTGCTGCTGTGGAGCAGTCTTCGTTTCGCCCGCTAGCGAAACCCACGT 86  
DB 1442 CTGGGCGAAGTCTCTGCTGCTATTTCCCGGCTCGACGGGAAACCCACGT 1501  
QY 87 CACCGGGGGAAGTGGCGGCCACACCTGTCGATTTGTTAGCTCTTCGACCGGCGC 146  
DB 1502 CACCGGGGGAAGTGGCGGCCCGCACCGCTGGCTTGTGCTCTTACACCGGCGC 1561  
QY 147 CAAGCAGAACCTGCTAGCTGATCAACACCAACGGCAGTGGCACCTCAATAGCAGCGCCCT 206  
DB 1562 CAAGCAGAACCTGCTAGCTGATCAACACCAACGGCAGTGGCACCTCAATAGCAGCGCCCT 1621  
QY 207 GAATCTGCAATGATAGCTCAACACCGGCTGGTGGCAGGCTTTCTATCACCACAGTT 266  
DB 1622 GAATCTGCAATGATAGCTTAAACCGGCTGGTGGCAGGCTTTCTATCACCACAAAT 1681  
QY 267 CAATCTCTCAGGCTGCTGAGAGCTAGCCAGCTGCCAGCCCTTACCGATTTTGACCA 326  
DB 1682 CAATCTCTCAGGCTGCTGAGAGTGGCCAGCTGCCAGCCCTTACCGATTTTGCCCA 1741  
QY 327 GGGCTGGGCTTATCAGTTATGCCAACCGAAGCGGCCCGGACGAGCCCTACTGCTG 386  
DB 1742 GGGCTGGGCTTATCAGTTATGCCAACCGAAGCGGCCCTCGACGAGCCCTACTGCTG 1801  
QY 387 GCATACCCCCCAAACTTCGGTATTTGTCGCCGGAAGAGTGTGTGGTCCGCTATA 446  
DB 1802 GCATACCCCTCAAGACCTTGTGGCAATTTGCCGCAAGAGCGTGTGGCCCGGTATA 1861  
QY 447 TTGCTTCACTCCAGCCCGTGTGGTGGGAACGACCGACAGTCCGGCGCGCCACCTTA 506  
DB 1862 TTGCTTCACTCCAGCCCGTGTGGTGGGAACGACCGACAGTCCGGCGCGCCACCTTA 1921  
QY 507 CAGCTGGGTGAAATGATACGGAOTCTTCGTCCTTAAACAATACAGGCCACCGCTGG 566  
DB 1922 CAGCTGGGTGCAATGATACGGAOTCTTCGTCCTTAAACAATACAGGCCACCGCTGG 1981  
QY 567 CAATGTTGCTGTTGATCTGATGATCACTCACTGATTTCAACCAAGTGTGGAGCGGC 626  
DB 1982 CAATGTTGCTGTTGATCTGATGATCACTCACTGATTTCAACCAAGTGTGGAGCGGC 2041  
QY 627 TCCCTGTGTCATCGGAGGGCGGGCAACACCTCTGCACTGCCCCACTGATTTGCTTCG 686  
DB 2042 CCGCTGTGTCATCGGAGGGGTGGCAACAACACTTGTCTGCCCCACTGATTTGTTCCG 2101  
QY 687 CAAGCATCCGGAGCGCCACATCTCTCGGTGGGCTCGGTCCTCGATCACACCGAGTG 746  
DB 2102 CAAGCATCCGGAGCGCCACATCTCTCGGTGGGCTCGGTCCTCGATCACACCGAGTG 2161  
QY 747 CCTGGTCGACTACCCGATAGGCTTTGGCAATATCTTGTACCATCACTACACCATATT 806  
DB

US-09-995-937-5  
Sequence 5, Application US/09995937  
Publication No. US20030028010A1  
GENERAL INFORMATION:  
APPLICANT: RICE, CHARLES et al.  
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOWELL & HAFERKAMP, L.C.  
STREET: 7733 FORSYTH BLVD., SUITE 1400  
CITY: ST. LOUIS  
STATE: MO  
COUNTRY: USA  
ZIP: 63105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/995,937  
FILING DATE: 28-No. US20030028010A1-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/034,756  
FILING DATE: 04-May-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: HOLLAND, DONALD R.  
REGISTRATION NUMBER: 35,197  
REFERENCE/DOCKET NUMBER: 6029-4831  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 314-727-5188  
TELEFAX: 314-727-6092  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12980 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-09-995-937-5

Query Match 46.8%; Score 745.2; DB 10; Length 12980;  
Best Local Similarity 90.1%; Pred. No. 7.5e-237;  
Matches 798; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 27 CTGTGCTGCTGCTGTGGAGCAGTCTTCGTTTCGCCCGCTAGCGAAACCCACGT 86  
DB 1442 CTGGGCGAAGTCTCTGCTGCTATTTCCCGGCTCGACGGGAAACCCACGT 1501  
QY 87 CACCGGGGGAAGTGGCGGCCACACCTGTCGATTTGTTAGCTCTTCGACCGGCGC 146  
DB 1502 CACCGGGGGAAGTGGCGGCCCGCACCGCTGGCTTGTGCTCTTACACCGGCGC 1561  
QY 147 CAAGCAGAACCTGCTAGCTGATCAACACCAACGGCAGTGGCACCTCAATAGCAGCGCCCT 206  
DB 1562 CAAGCAGAACCTGCTAGCTGATCAACACCAACGGCAGTGGCACCTCAATAGCAGCGCCCT 1621  
QY 207 GAATCTGCAATGATAGCTCAACACCGGCTGGTGGCAGGCTTTCTATCACCACAGTT 266  
DB 1622 GAATCTGCAATGATAGCTTAAACCGGCTGGTGGCAGGCTTTCTATCACCACAAAT 1681  
QY 267 CAATCTCTCAGGCTGCTGAGAGCTAGCCAGCTGCCAGCCCTTACCGATTTTGACCA 326  
DB 1682 CAATCTCTCAGGCTGCTGAGAGTGGCCAGCTGCCAGCCCTTACCGATTTTGCCCA 1741  
QY 327 GGGCTGGGCTTATCAGTTATGCCAACCGAAGCGGCCCGGACGAGCCCTACTGCTG 386  
DB 1742 GGGCTGGGCTTATCAGTTATGCCAACCGAAGCGGCCCTCGACGAGCCCTACTGCTG 1801  
QY 387 GCATACCCCCCAAACTTCGGTATTTGTCGCCGGAAGAGTGTGTGGTCCGCTATA 446  
DB 1802 GCATACCCCTCAAGACCTTGTGGCAATTTGCCGCAAGAGCGTGTGGCCCGGTATA 1861  
QY 447 TTGCTTCACTCCAGCCCGTGTGGTGGGAACGACCGACAGTCCGGCGCGCCACCTTA 506  
DB 1862 TTGCTTCACTCCAGCCCGTGTGGTGGGAACGACCGACAGTCCGGCGCGCCACCTTA 1921  
QY 507 CAGCTGGGTGAAATGATACGGAOTCTTCGTCCTTAAACAATACAGGCCACCGCTGG 566  
DB 1922 CAGCTGGGTGCAATGATACGGAOTCTTCGTCCTTAAACAATACAGGCCACCGCTGG 1981  
QY 567 CAATGTTGCTGTTGATCTGATGATCACTCACTGATTTCAACCAAGTGTGGAGCGGC 626  
DB 1982 CAATGTTGCTGTTGATCTGATGATCACTCACTGATTTCAACCAAGTGTGGAGCGGC 2041  
QY 627 TCCCTGTGTCATCGGAGGGCGGGCAACACCTCTGCACTGCCCCACTGATTTGCTTCG 686  
DB 2042 CCGCTGTGTCATCGGAGGGGTGGCAACAACACTTGTCTGCCCCACTGATTTGTTCCG 2101  
QY 687 CAAGCATCCGGAGCGCCACATCTCTCGGTGGGCTCGGTCCTCGATCACACCGAGTG 746  
DB 2102 CAAGCATCCGGAGCGCCACATCTCTCGGTGGGCTCGGTCCTCGATCACACCGAGTG 2161  
QY 747 CCTGGTCGACTACCCGATAGGCTTTGGCAATATCTTGTACCATCACTACACCATATT 806  
DB

```

QY 207 GAAGTCAATGATAGCCTCAACACCGGCTGGTTGGCAGGCGCTTTCTATCACCACAAGTT 266
Db 1622 GAAGTCAATGAAAGCCCTTAACACCGGCTGGTTAGCAGGCGCTCTCTATCAGCACAAAT 1681
QY 267 CAATCTTTAGGCTGCTCTGAGAGGTAGCCAGCTGCGGACCCCTTACCGATTTTGACCA 326
Db 1682 CAATCTTTAGGCTGCTCTGAGAGGTGGCCAGCTGCGGACCCCTTACCGATTTTGCCCA 1741
QY 327 GGCTGGGGCCCTATCAGTTTATCCCAAGGAAGCGGCCCGACAGCGCCCTACTGCTG 386
Db 1742 GGCTGGGGTCTTATCAGTTTATCCCAAGGAAGCGGCCCTGAGAAAGCGCCCTACTGCTG 1801
QY 387 GCATCAACCCCAACCTTTGCGGTATTTGCGCCGGAAGAGTGTGTGTTGTCGGTATA 446
Db 1802 GCATCAACCTTCAAGACCTTTGCGCATTTGCGCCGGAAGAGGTGTGTGTCGGTATA 1861
QY 447 TTGCTTCACTCCAGCGCCGCTGTGTGGAAGACGACGACAGTGTGCGGCGGCCACCTA 506
Db 1862 TTGCTTCACTCCAGCGCCGCTGTGTGGAAGACGACGACAGTGTGCGGCGGCCCTACTA 1921
QY 507 CAGCTGGGTGAAATGATACGACGCTCTTCGTCCTTAAACAATACCAGGCGCCGCTGGG 566
Db 1922 CAGCTGGGTGCAATGATACGAGTGTCTTCGTCCTTAAACAATACCAGGCGCCGCTGGG 1981
QY 567 CAATTGGTTGGTTGTTACTGTGAATCAACTCAACTGGATTTCACCAAGTGTGGGAGCGCC 626
Db 1982 CAATTGGTTGGTTGTTACTGTGAATCAACTCAACTGGATTTCACCAAGTGTGGGAGCGCC 2041
QY 627 TCCTTGTGTCTATCGGAGGGGGGCAACAACACCTCTGACCTGCCCTGATTTGCTTCG 686
Db 2042 CCCTTGTGTCTATCGGAGGGGGGCAACAACACCTTGTCTGCCCCCTGATTTGTTCCG 2101
QY 687 CAAGCATCCGAGCGGCACATACCTCTCGTGGCGCTCCGTCCTGTGATCACACCCAGTG 746
Db 2102 CAAGCATCCGAGCGGCACATACCTCTCGTGGCGCTCCGTCCTGTGATCACACCCAGTG 2161
QY 747 CTGTGTCGACTACCCGTTAGGCTTTGGCATTTATCTTGTACCATCAACTACACCATATT 806
Db 2162 CATGTCGACTACCCGTTATAGGCTTTGGCATTTATCTTGTACCATCAANTACACCATATT 2221
QY 807 TAAATCAGATGATCAGTGGAGGGGTGGAACACAGCTGGAAGCTGCAATCTGCAATCTGAC 866
Db 2222 CAAAGTCAGATGATCAGTGGAGGGGTGGAACACAGCTGGAAGCTGCAATCTGCAATCTGAC 2281
QY 867 GCGGGCGGAACGTTGGATCTGGAATAGGAGACAGGTCGGAGATC 912
Db 2282 GCGGGCGGAACGTTGGATCTGGAATAGGAGACAGGTCGGAGATC 2327

```

RESULT 12

```

US-09-917-563-5
; Sequence 5, Application US/09917563
; Publication No. US2003007080A1
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/917,563

```

```

; FILING DATE: 27-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/238,076
; FILING DATE: 26-JAN-1999
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
; TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-917-563-5

```

```

Query Match 46.8%; Score 745.2; DB 10; Length 12980;
Best Local Similarity 90.1%; Pred. No. 7.5e-237;
Matches 798; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 27 CTGTGTGCTGCTGTGTGGAGAGTCTTCTGTTGCGCCAGCGGTAGGAAACCCACGT 86
Db 1442 CTGGCGCAAGGCTCTGGTAGTCTGTCTATTTGCGCGCTCGACGCGAAACCCACGT 1501
QY 87 CACCGGGGAAGTGGCGGCACACTGTGTCGATTTGTAGCTCTCGCACCGCGCC 146
Db 1502 CACCGGGGAAGTGGCGGCACCAACCGCTGTGTGTGCTCTCTTACACGCGC 1561
QY 147 CAAGCAGAACGTCAGCTGATCAACACCAACCGGAGTTGGCACCTCAATAGCACGGCCCT 206
Db 1562 CAAGCAGAACATCCAACTGATCAACCAACCGGAGTTGGCACATCAATAGCACGGCCCT 1621
QY 207 GAATCTGAATGATAGCTCAACACCGGCTGTGTGGAGGCTTTTCTATCACCAAGTT 266
Db 1622 GAATCTGAATGAAAGCTTAAACCGGCTGTGTAGCAGGCTCTTCTATCAGCACAAAT 1681
QY 267 CAACTCTTCAGGCTGCTCTGAGAGCTAGCCAGCTGCGGACCCCTTACCGATTTTGACCA 326
Db 1682 CAACTCTTCAGGCTGCTCTGAGAGGTGGCCAGCTGCGGACCCCTTACCGATTTTGCCCA 1741
QY 327 GGGCTGGGGCCCTATCAGTTATGCCAAGCGGCCCGCCAGCGCCCTACTGCTG 386
Db 1742 GGGCTGGGGTCTTATCAGTTATGCCAAGCGGCCCTCGACGAAAGCGCCCTACTGCTG 1801
QY 387 GCATTAACCCCAACCTTTGCGGTATTTGTCGCCGGAAGAGTGTGTGTGTCGGTATA 446
Db 1802 GCATTAACCTTCAAGACCTTTGCGCATTTGTCGCCCAAGAGCGTGTGTGGCCGCTATA 1861
QY 447 TTGCTTCACTCCAGCGCCGCTGTGTGGAGACGACGACAGTGTGGGCGCGCCACCTA 506
Db 1862 TTGCTTCACTCCAGCGCCGCTGTGTGGAGACGACGACAGTGTGGGCGCGCCCTACTA 1921
QY 507 CAGCTGGGTGAAATGATACGAGCTCTTCTGTCCTTAAACAATACCAGGCGCCGCTGGG 566
Db 1922 CAGCTGGGTGCAATGATACGAGTGTCTTCGTCCTTAAACAATACCAGGCGCCGCTGGG 1981
QY 567 CAATTGGTTGGTTGTTACTGTGAATCAACTCAACTGGATTTCACCAAGTGTGGGAGCGCC 626
Db 1982 CAATTGGTTGGTTGTTACTGTGAATCAACTCAACTGGATTTCACCAAGTGTGGGAGCGCC 2041
QY 627 TCCTTGTGTCTATCGGAGGGGGGCAACAACACCTCTGACCTGCCCTGATTTGCTTCG 686
Db 2042 CCCTTGTGTCTATCGGAGGGGGGCAACAACACCTTGTCTGCCCCCTGATTTGTTCCG 2101
QY 687 CAAGCATCCGAGCGGCACATACCTCTCGTGGCGCTCCGTCCTGTGATCACACCCAGTG 746

```









Search completed: April 15, 2005, 04:52:24  
Job time : 1000 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2005, 21:43:49 ; Search time 5460 Seconds  
(without alignments)  
11105.564 Million cell updates/sec

Title: US-10-715-665-6\_COPY\_1992\_3584

Perfect score: 1593

Sequence: 1 atgtagcaatgaagagag.....ttgtctctgggtatacatt 1593

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.\*

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_btc.\*
- 4: gb\_est3.\*
- 5: gb\_est4.\*
- 6: gb\_est5.\*
- 7: gb\_est6.\*
- 8: gb\_gss1.\*
- 9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	4.3	290	5 BQ678991	BQ678991 AGENCOURT
2	69	4.3	300	1 AU100324	AU100324 AU100324
3	69	4.3	378	1 AA317921	AA317921 EST20057
4	69	4.3	393	6 CB139652	CB139652 K-EST0192
5	69	4.3	469	7 CN268878	CN268878 170004247
6	69	4.3	495	2 AW2339072	AW2339072 xb36b11.y
7	69	4.3	500	6 CB129127	CB129127 K-EST0178
8	69	4.3	511	2 BE389162	BE389162 601285961
9	69	4.3	531	6 CD672904	CD672904 fg17f07.y
10	69	4.3	571	1 AU134450	AU134450 AU134450
11	69	4.3	578	5 BP366704	BP366704 BP366704
12	69	4.3	580	1 AU138842	AU138842 AU138842
13	69	4.3	581	7 CN481173	CN481173 hw03h06.y
14	69	4.3	584	5 BP259505	BP259505 BP259505
15	69	4.3	593	6 CB131100	CB131100 K-EST0181
16	69	4.3	594	7 CN483734	CN483734 hw36f09.y
17	69	4.3	612	2 BE386111	BE386111 601277221
18	69	4.3	616	7 CN268875	CN268875 170005325
19	69	4.3	620	7 CF132460	CF132460 UI-HF-FQ0
20	69	4.3	624	2 BE390617	BE390617 601284976
21	69	4.3	632	1 AU124041	AU124041 AU124041
22	69	4.3	636	7 CF125573	CF125573 UI-HF-EL0
23	69	4.3	637	1 AU134370	AU134370 AU134370
24	69	4.3	641	7 CV027090	CV027090 5255 Full

25	69	4.3	646	1 AU134236	AU134236 AU134236
26	69	4.3	646	4 BG477928	BG477928 602522906
27	69	4.3	650	7 CN268879	CN268879 170005319
28	69	4.3	658	2 BE384999	BE384999 601276871
29	69	4.3	680	7 CF264856	CF264856 AGENCOURT
30	69	4.3	680	7 CN268877	CN268877 170006001
31	69	4.3	686	7 CF125651	CF125651 UI-HF-EL0
32	69	4.3	688	7 CO245698	CO245698 AGENCOURT
33	69	4.3	694	1 AL524377	AL524377 AL524377
34	69	4.3	695	1 AU134787	AU134787 AU134787
35	69	4.3	703	2 BE390092	BE390092 601285720
36	69	4.3	706	1 AL036128	AL036128 DKF2p5648
37	69	4.3	710	2 BE261317	BE261317 601148758
38	69	4.3	711	4 BG476511	BG476511 602522049
39	69	4.3	711	4 BG763226	BG763226 602735258
40	69	4.3	716	2 BE408331	BE408331 601302704
41	69	4.3	720	2 BF689648	BF689648 602187048
42	69	4.3	723	7 CO245365	CO245365 AGENCOURT
43	69	4.3	740	4 BG768016	BG768016 602743779
44	69	4.3	755	4 BG761429	BG761429 602718630
45	69	4.3	761	4 BI771956	BI771956 603058981

ALIGNMENTS

BQ678991 290 bp mRNA linear EST 15-JUL-2002  
AGENCOURT\_8281562 NIH\_MGC\_112 Homo sapiens cDNA clone IMAGE:6262005  
5', mRNA sequence.  
ACCESSION BQ678991.1 GI:21791670  
VERSION BQ678991.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 290)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-rc@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM2423 row: j column: 22  
High quality sequence stop: 289.  
Location/Qualifiers  
1. 290  
/organism="Homo sapiens"  
/mol type="mRNA"  
/db xref="taxon:9606"  
/clone="IMAGE:6262005"  
/tissue\_type="melanotic melanoma, cell line"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH MGC 112"  
/notes="Organ: sKin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 4.3%; Score 69; DB 5; Length 290;



SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 393)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsaung@mail.kribb.re.kr  
Plate: 19 Row: B Column: 03  
High quality sequence stop: 393.

FEATURES  
source  
1. .393  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L15CKK1-19-B03"  
/sex="M"  
/cell\_line="CK-K1"  
/lab\_hosts="Top10P"  
/clone\_lib="L15CKK1"  
/note="Organ: Liver; Vector: pCNS-D2; Site\_1: EcoRI; Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-tailed vector priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN  
Query Match 4.3%; Score 69; DB 6; Length 393;  
Best Local Similarity 100.0%; Pred. No. 4e-09;  
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGTCGAGCAGTCTTCGTT 60  
|||||  
Db 101 ATCGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGTCGAGCAGTCTTCGTT 160  
|||||

QY 61 TCGCCCCAGC 69  
|||||  
Db 161 TCGCCCCAGC 169

RESULT 5  
LOCUS CN268878  
DEFINITION 17000424723882 GRN\_ES Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION CN268878  
VERSION CN268878.1 GI:47285292  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 469)  
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,

Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.  
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 469 Std Error: 0.00.

FEATURES  
source  
1. .469  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, cell lines H1, H7, and H9"  
/clone\_lib="GRN\_ES"  
/note="oligo dt primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN  
Query Match 4.3%; Score 69; DB 7; Length 469;  
Best Local Similarity 100.0%; Pred. No. 4.2e-09;  
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGTCGAGCAGTCTTCGTT 60  
|||||  
Db 86 ATCGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGTGTCGAGCAGTCTTCGTT 145  
|||||

QY 61 TCGCCCCAGC 69  
|||||  
Db 146 TCGCCCCAGC 154

RESULT 6  
LOCUS AW239072  
DEFINITION AW239072 495 bp mRNA linear EST 13-DEC-1999  
similar to gb:M15518 TISSUE PLASMINOGEN ACTIVATOR PRECURSOR (HUMAN); mRNA sequence.

ACCESSION AW239072  
VERSION AW239072.1 GI:6571462  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 495)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-f@mail.nih.gov  
Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Seq primer: -40RP from Gibco  
High quality sequence stop: 456.

FEATURES  
source  
1. .495  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

obtained cDNA vectors were used for transformation of competent cells *E. coli* Top10F<sup>+</sup> by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

[illegible]

RESULT	8
LOCUS	BE389162
DEFINITION	01285961F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3607767 5', mRNA linear EST 21-JUL-2000
ACCESSION	BE389162
VERSION	BE389162
KEYWORDS	BE389162.1 GI:9334527
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ 1 (bases 1 to 511)
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: craebbs@mail.nih.gov
COMMENT	

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLCM259 row: i column: 16  
High quality sequence stop: 510.

```

FEATURES
source
Location/Qualifiers
1. 5311
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3607767"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_44"
/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN					
Query Match            4.3%; Score 69; DB 2; Length 511;					
Best Local Similarity 100.0%; Pred. No. 4.3e-09;					
Matches	69;	Conservative	0;	Mismatches	0;
				Indels	0;
Gaps	0;				
Qy	1	ATGGATGCAATGAAGACAGAGGCCTCGTGTGTGCTGCTGTGTGGAGCAGTCCTTCGTT	60		
Dd	58	ATGGATGCATGAAGACAGAGGCCTCGTGTGTGCTGCTGTGTGGAGCAGTCCTTCGTT	117		
Qy	61	TCGCCCCAGC	69		

```

Db      118 TCGCCGAGC 126

RESULT 9
CD672904      531 bp      mRNA      linear      EST 24-JUN-2003
LOCUS      fg17f07.y1 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
DEFINITION      fg17f07 5', mRNA sequence.
ACCESSION      CD672904
VERSION      CD672904
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 531)
AUTHORS      Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
              Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE      Expressed sequence tag analysis of adult human iris for the NEIBank
              Project: steroid-response factors and similarities with retinal
              pigment epithelium
JOURNAL      Mol. Vis. 8 (4), 185-195 (2002)
MEDLINE      22103462
PubMed      12107412
COMMENT      Contact: Wistow G
              Section on Molecular Structure and Function
              National Eye Institute
              6/331, NIH, Bethesda, MD 20892-2740, USA
              Tel: 301 402 3452
              Fax: 301 496 0078
              Email: graeme@helix.nih.gov
              Plate: 17 row: f column: 07
              Seq primer: M13Rpl reverse primer (ABT).
              Location/Qualifiers
FEATURES      source
              1..531
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="fg17f07"
                /tissue_type="Iris"
                /dev_stage="Adult"
                /lab_host="EMDH108"
                /clone_lib="Human Iris cDNA (Normalized): fg"
                /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
              library (bx) was normalized by self-subtraction. One
              portion of double stranded plasmid DNA representing the
              library was linearized by NotI. This NotI digested library
              was used as a template for biotinylated RNA synthesis
              using SP6 RNA polymerase. Another portion of the double
              stranded plasmid library was converted to single-stranded
              circles in vitro using Gene II and Exonuclease III (Life
              Technologies). Single-stranded DNA (1 mg) was hybridized
              (60t 500) with 41 mg of Bio-RNA and vector blocking
              oligonucleotides. The hybridized Bio-RNA/ss-circles were
              removed by streptavidin:phenol extraction. EST analysis
              was performed on the library at the NIH Intramural
              Sequencing Center (NISC)."
ORIGIN
Query Match      4.3%; Score 69; DB 6; Length 531;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCAGTCTTCGTT 60
        |||||||
Db      113 ATGGATGCAATGAAGAGAGGGCTCTGCTGTGCTGCTGCTGTGGAGCAGTCTTCGTT 172
        |||||||

Qy      61 TCGCCGAGC 69
        |||||||
Db      173 TCGCCGAGC 181
        |||||||

RESULT 11
CD672904      578 bp      mRNA      linear      EST 17-SEP-2004
LOCUS      BP366704 Sugano cDNA library, thymus Homo sapiens cDNA clone
DEFINITION      BP366704.1 GI:52296954
ACCESSION      BP366704
VERSION      BP366704
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      1 (bases 1 to 578)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              Suzuki,Y., Yamaguchi,R., Shirota,M., Sakakibara,Y., Chiba,J.,
              Mizushima-Sugano,J., Nakai,K. and Sugano,S.
              Sequence comparison of human and mouse genes reveals a homologous
              block structure in the promoter regions
              Genome Res. 14 (9), 1711-1718 (2004)
              Contact: Yutaka Suzuki
              Department of Virology
TITLE
JOURNAL
COMMENT
  
```

Institute of Medical Science, University of Tokyo  
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
Email: ysuzuki@ims.u-tokyo.ac.jp.

## FEATURES

source

```

source
1..578
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TMS01388"
/tissue_type="thymus"
/clone_lib="Sugano cdna library, thymus"

ORIGIN
      4.3%;   Score 69;   DB 5;   Length 578;
Query Match
Best Local Similarity 100.0%;   Pred. No. 4.5e-09;
Matches 69;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0

Qy      1  ATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGAGGACACTCTTCGTT 60
        |||||
Db      207 ATGGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGAGGACAGTCTTCGTT 266

Qy      61  TCGCCCCAGC 69
        |||||
Db      267 TCGCCCCAGC 275

```

## ORIGIN

Query Match	4.38	Score 69	DB 5	Length 578
Best Local Similarity	100.0%	Prod. No. 4.5e-09		
Matches 69	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	ATGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGTGCTGTGCTGTGAGCAGCTTCGTT	60	
Db	207	ATGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGTGCTGTGCTGTGAGCAGCTTCGTT	256	
Qy	61	TCGCCACG	69	
Db	267	TCGCCACG	275	

RESULT 12	
AUI138842	
LOCUS	580 bp mRNA linear EST 02-AUG-2002
DEFINITION	AUI138842 PLACE1 Homo sapiens cDNA clone PLACE1009403 5', mRNA sequence.
ACCESSION	AUI138842
VERSION	AUI138842.1 GI:11000363
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 580)
AUTHORS	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yanamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.

**TITLE**

Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

## FEATURES

```

1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1009403"
/tissue_type="placenta"
/clone_lib="pLACel"
/note="Vector: pME18SFL3"

```

## ORIGIN

	Query Match	4.3%	Score 69;	DB 1;	Length 580;
	Best Local Similarity	100.0%;	Pred. No. 4.5e-09;		
	Matches 69;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGATGCAATCAAGAGAGGGCTTCGCTGTGTGCTGCTGTGTGGAGCAGCTTTCGTT	60		
Dp	98	ATGGATGCAATCAAGAGAGGGCTTCGCTGTGTGCTGCTGTGTGGAGCAGCTTTCGTT	157		

Qy	61	TCGCCGAGC 69       
Dd	158	TCGCCGAGC 166
RESULT 13		
CN481173	CN481173	581 bp mRNA linear EST 26-APR-2004
LOCUS	hw03h06.y1	Human primary human ocular pericytes. Unamplified (hw)
DEFINITION	Homo sapiens cDNA clone hw03h06 5', mRNA sequence.	
ACCESSION	CN481173	
VERSION	CN481173.1	GI:46562677
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 581) Tsai J.Y. and Wistow,G.	
AUTHORS	Expressed sequence tag analysis of cultured primary human ocular pericytes	
TITLE	Unpublished (2004)	
JOURNAL	Contact: Wistow G	
COMMENT	Section on Molecular Structure and Function National Eye Institute 6/331, NIH, Bethesda, MD 20892-2740, USA Tel: 301 402 3452 Fax: 301 496 0078 Email: graeme@helix.nih.gov plate: 03 row: h column: 06 Seg primer: M13Rp1 reverse primer (ABI).	

## FEATURES

```

1. .591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hw03h06"
/cell_type="pericytes"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human primary human ocular pericytes.
Unamplified (hw)"
name="Organ; Eye; Vector: pSport1; RNA was extracted from
primary human pericytes in culture. A directionally cloned
cDNA library in the pSPORT1 vector (Invitrogen) was
constructed at Bioserve Biotechnology (Laurel MD)
essentially following the protocols of the SuperScript
Plasmid System full details of which are contained in the
manufacturer's instruction manual
(http://www.lifetech.com/). First strand synthesis was
carried out using a Not I primer-adaptor
[5'-pGACTAGTTCAGTCGAGCGGCCCT(1)15-3']. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed at
the NIH Intramural Sequencing Center (NISC)."
```

## ORIGIN

Query Match	4.3%;	Score 69;	DB 7;	Length 581;
Best Local Similarity	100.0%;	Pred. No. 4.5e-09;		
Matches 69;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	ATGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGCTGTGAGGAGCTCTTCGTT	60	
Db	64	ATGATGCAATGAAGAGAGGGCTCTGCTGTGTGCTGCTGTGCTGTGAGGAGCTCTTCGTT	123	
Qy	61	TCGCCCCAGC	69	
Db	124	TCGCCCCAGC	132	
RESULT 14				
BP259505				
LOCUS	BP259505	584 bp	mrna	linear
				EST 16-SEP-2000





**This Page Blank (uspto)**